

# BGP in 2013 (and a bit of 2014)

Geoff Huston  
APNIC

RiPE 68

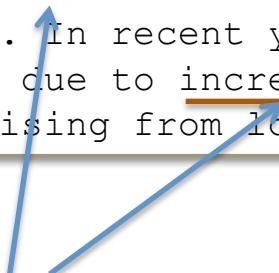
## “Conventional “wisdom” about routing:

“The rapid and sustained growth of the Internet over the past several decades has resulted in large state requirements for IP routers. In recent years, these requirements are continuing to worsen, due to increased deaggregation (advertising more specific routes) arising from load balancing and security concerns..”

Quoted from a 2012 research paper on routing

## “Conventional “wisdom” about routing:

“The rapid and sustained growth of the Internet over the past several decades has resulted in large state requirements for IP routers. In recent years, these requirements are continuing to worsen, due to increased deaggregation (advertising more specific routes) arising from load balancing and security concerns..”



quote from a 2012

routing

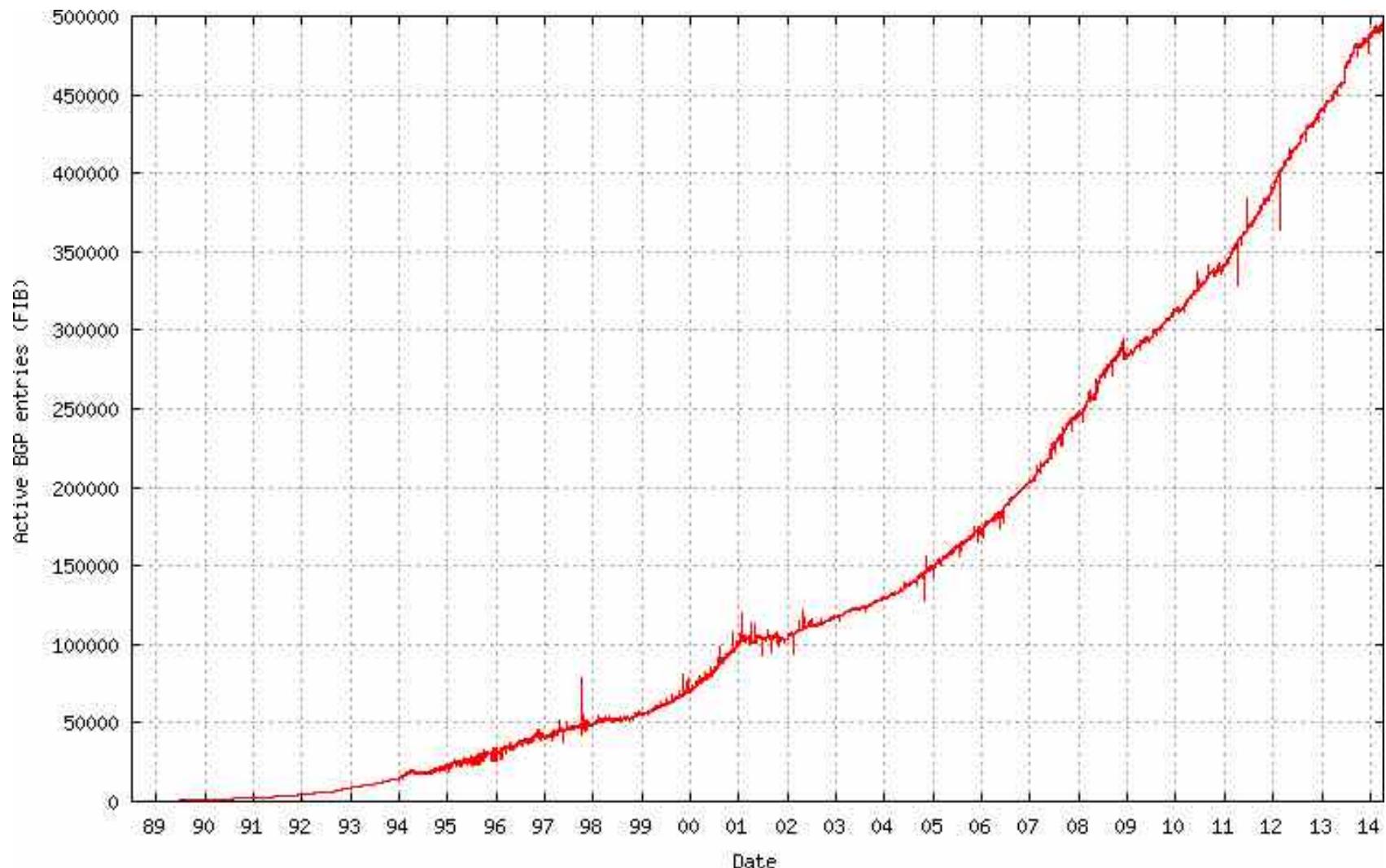
is this really true, or do we accept it as true without actually looking at the real behaviours of the internet's routing system???

In this presentation we'll explore the space of inter-domain routing and look at

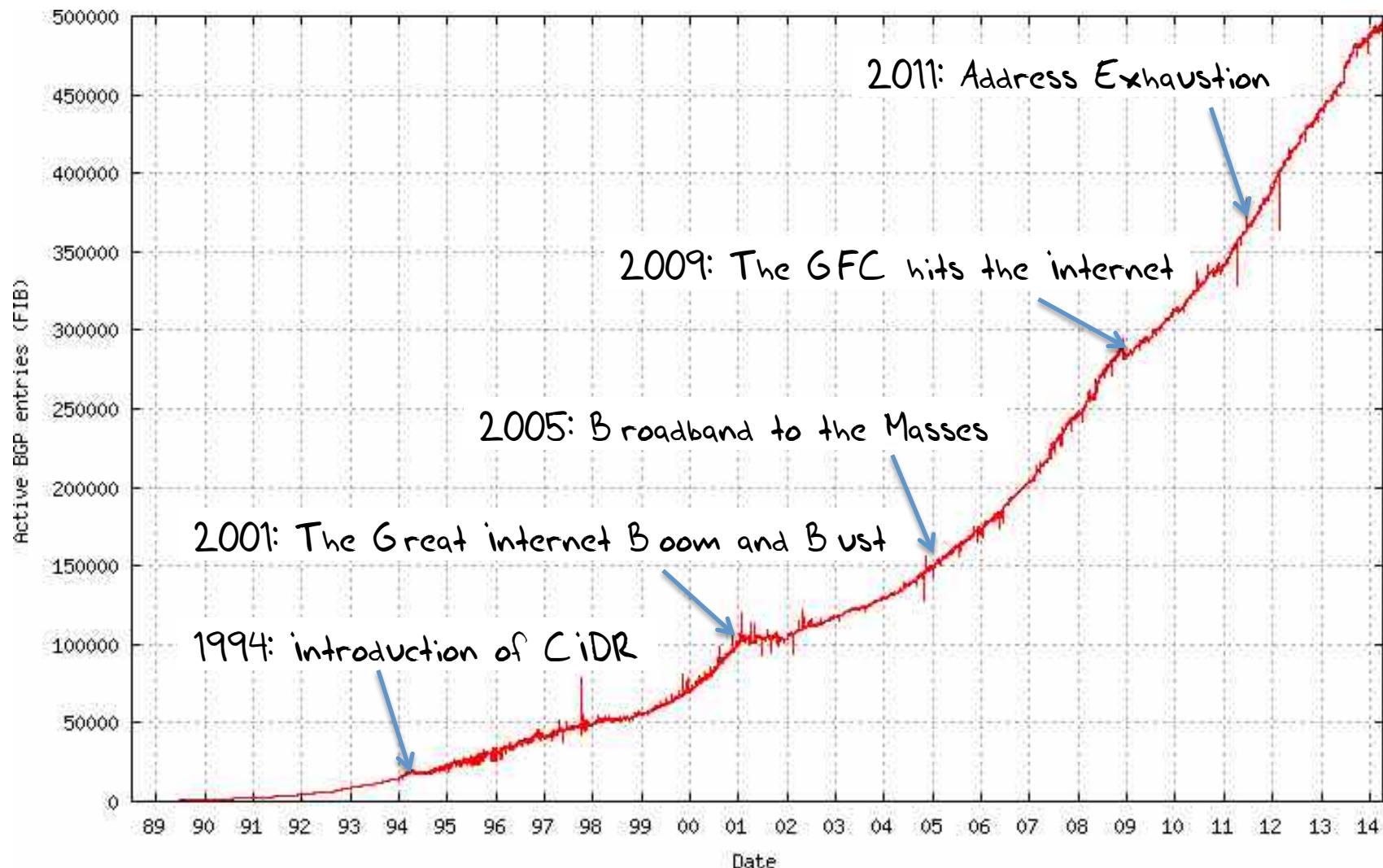
- the growth of the eBGP routing table over time and some projections for future growth
- the extent to which more specifics are dominating routing table growth ... or not

# What we saw

# The Big Picture of the v4 Routing Table

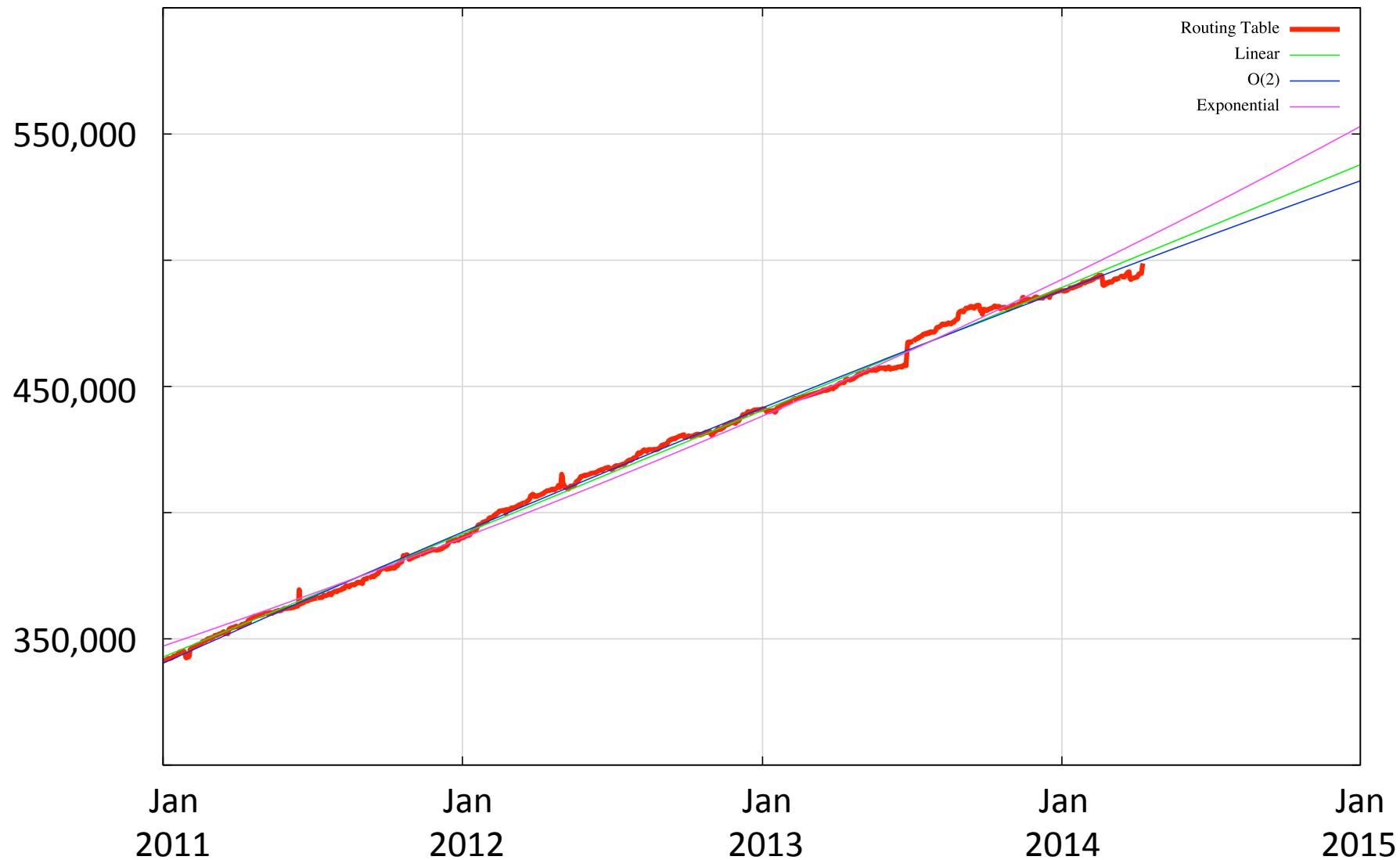


# The Big Picture of the v4 Routing Table

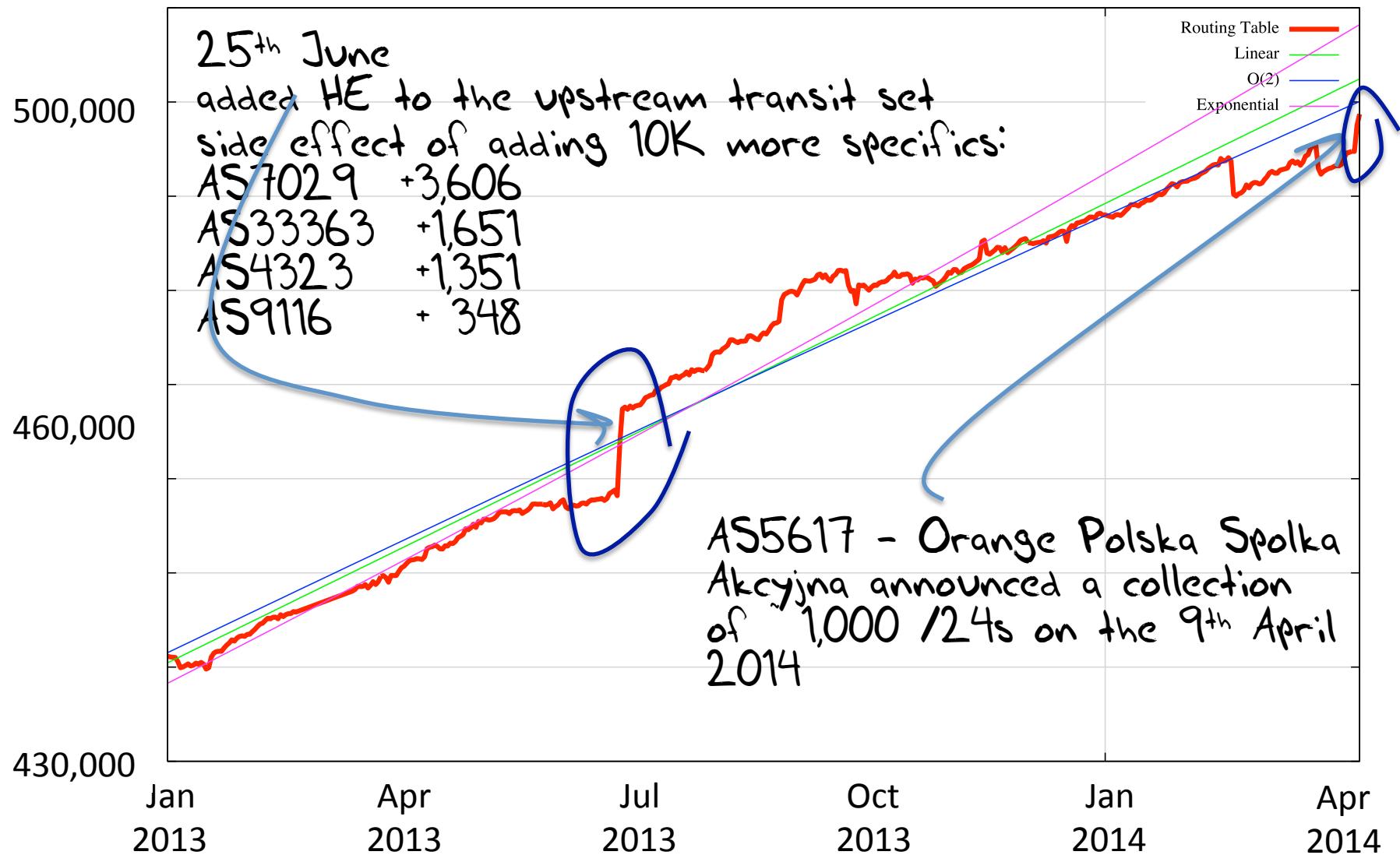


Lets look at the recent past in a little more detail...

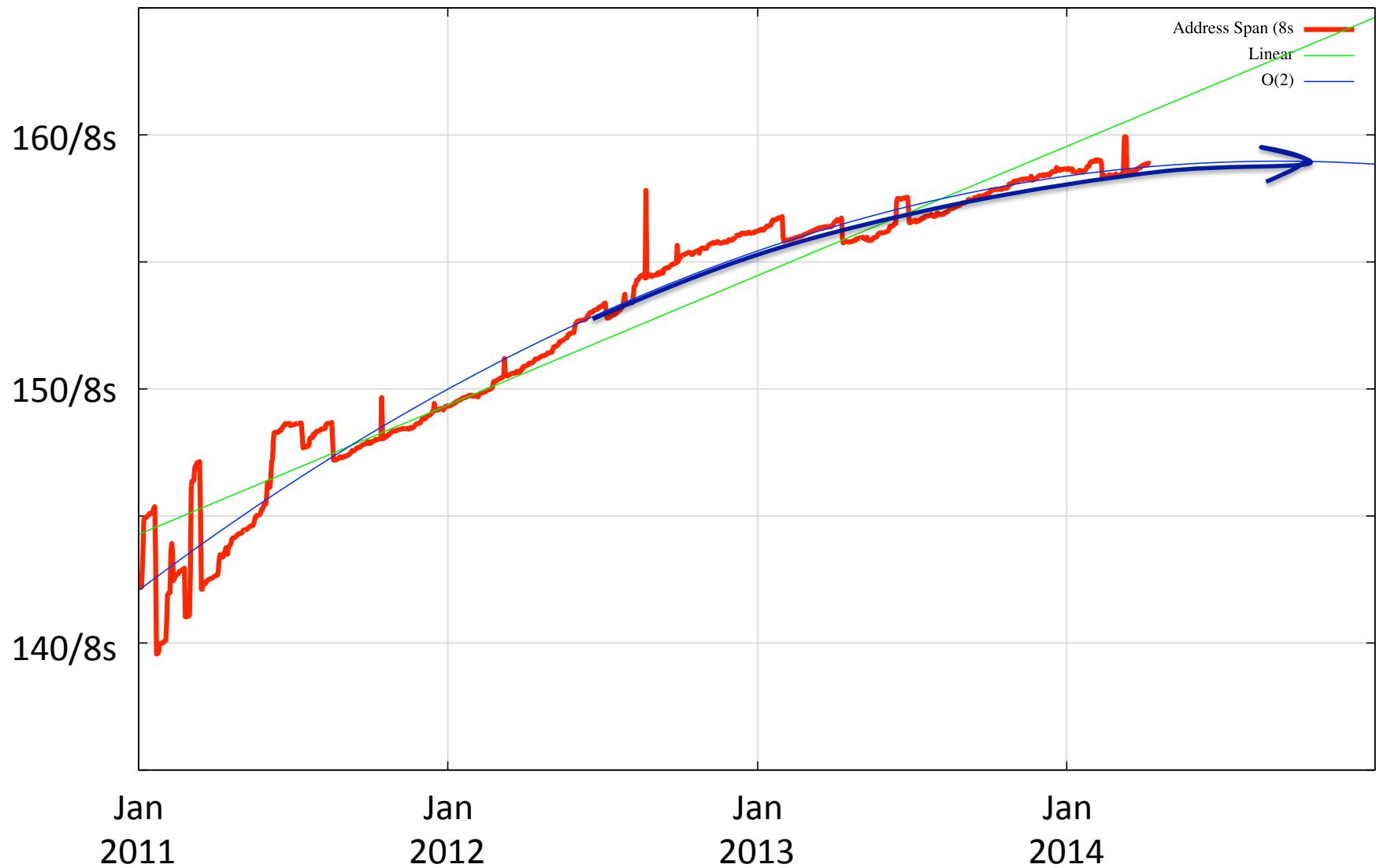
# IPv4 BGP Prefix Count 2011 - 2014



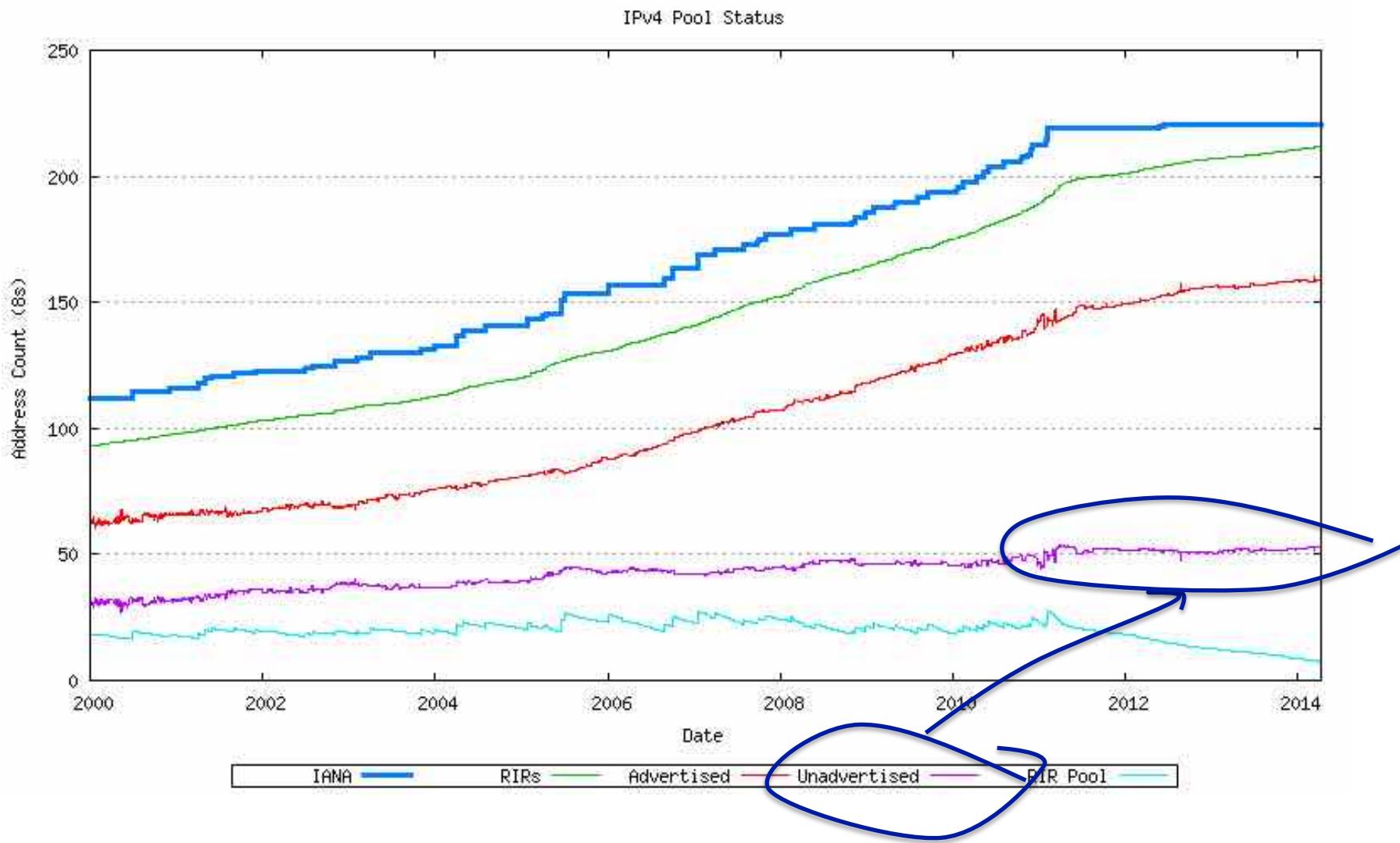
# IPv4 BGP Prefix Count 2013 - 2014



# IPv4 Routed Address Span: 2011 - 2013



# IPv4 Address Pool



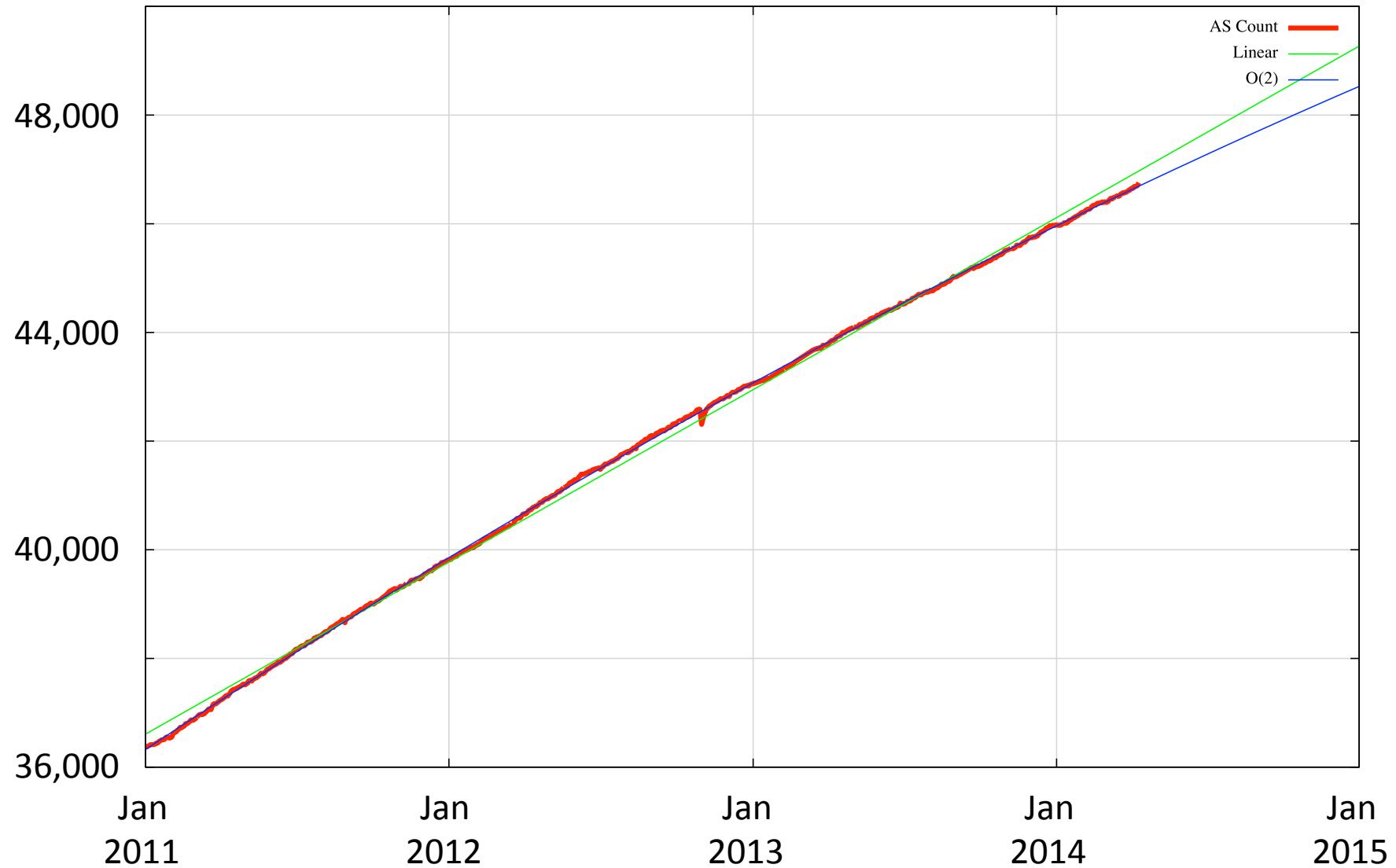
# The Unadvertised IPv4 Address Pool

Appears to be relatively static in size since early 2011 at some 50 /8s, or 20% of the IPv4 global unicast space.

So far whatever address transfers and trading may have taken place has not had a significant impact on the levels of unadvertised IPv4 address space. This may be due to:

- the last /8 policies of APNIC and the RIPE NCC, or
- the remaining address pools in ARIN and LACNIC, or
- a reluctance on the part of these address holders to trade their addresses at this point in time

# IPv4 Routed AS Count



# IPv4 2013 BGP Vital Statistics

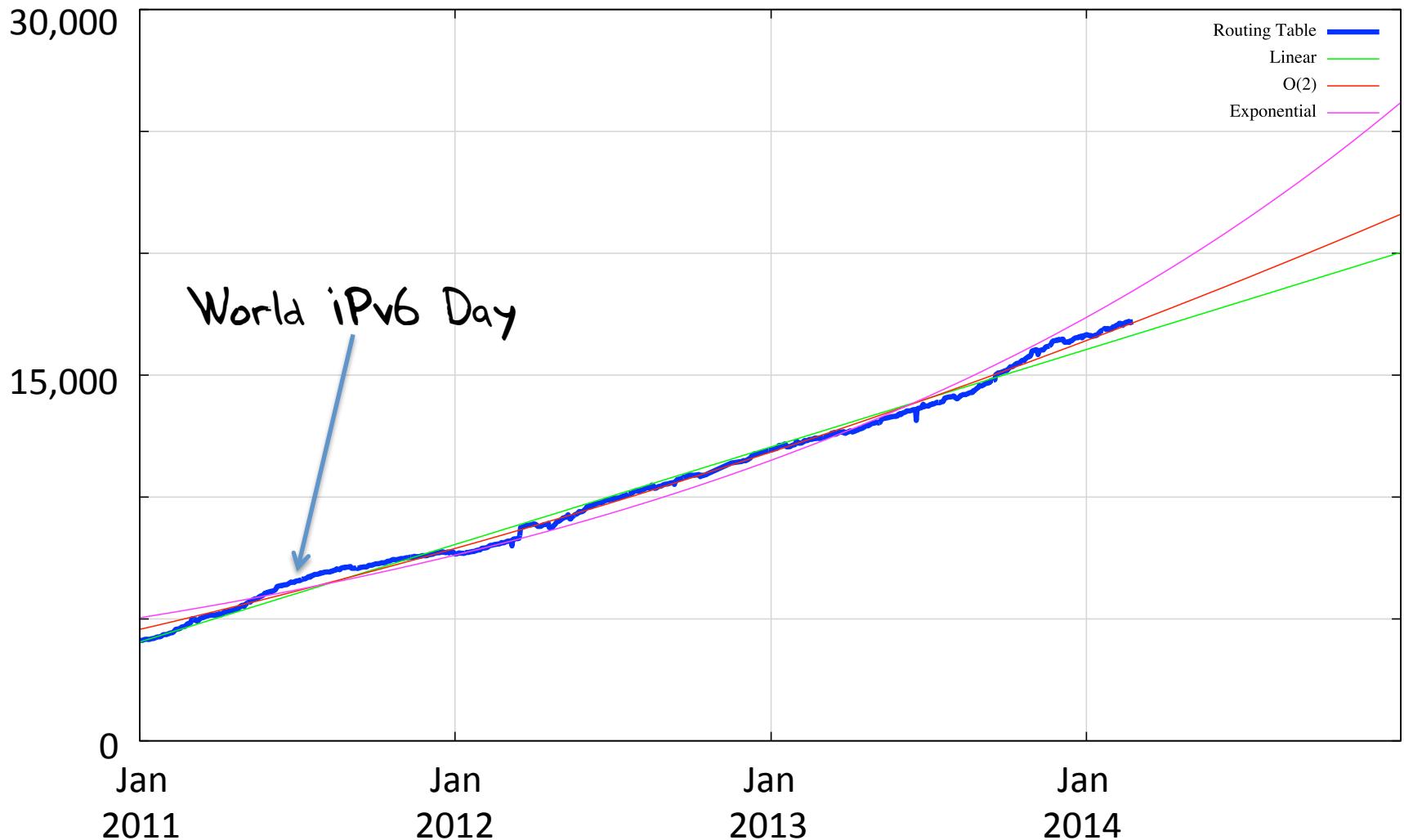
	Jan-13	Jan-14	
<b>Prefix Count</b>	440,000	488,000	+11%
Roots	216,000	237,000	+10%
More Specifics	224,000	251,000	+12%
<b>Address Span</b>	156/8s	159/8s	+ 2%
<b>AS Count</b>	43,000	46,000	+ 7%
Transit	6,100	6,600	+ 8%
Stub	36,900	39,400	+ 7%

# IPv4 in 2013 – Growth is Slowing

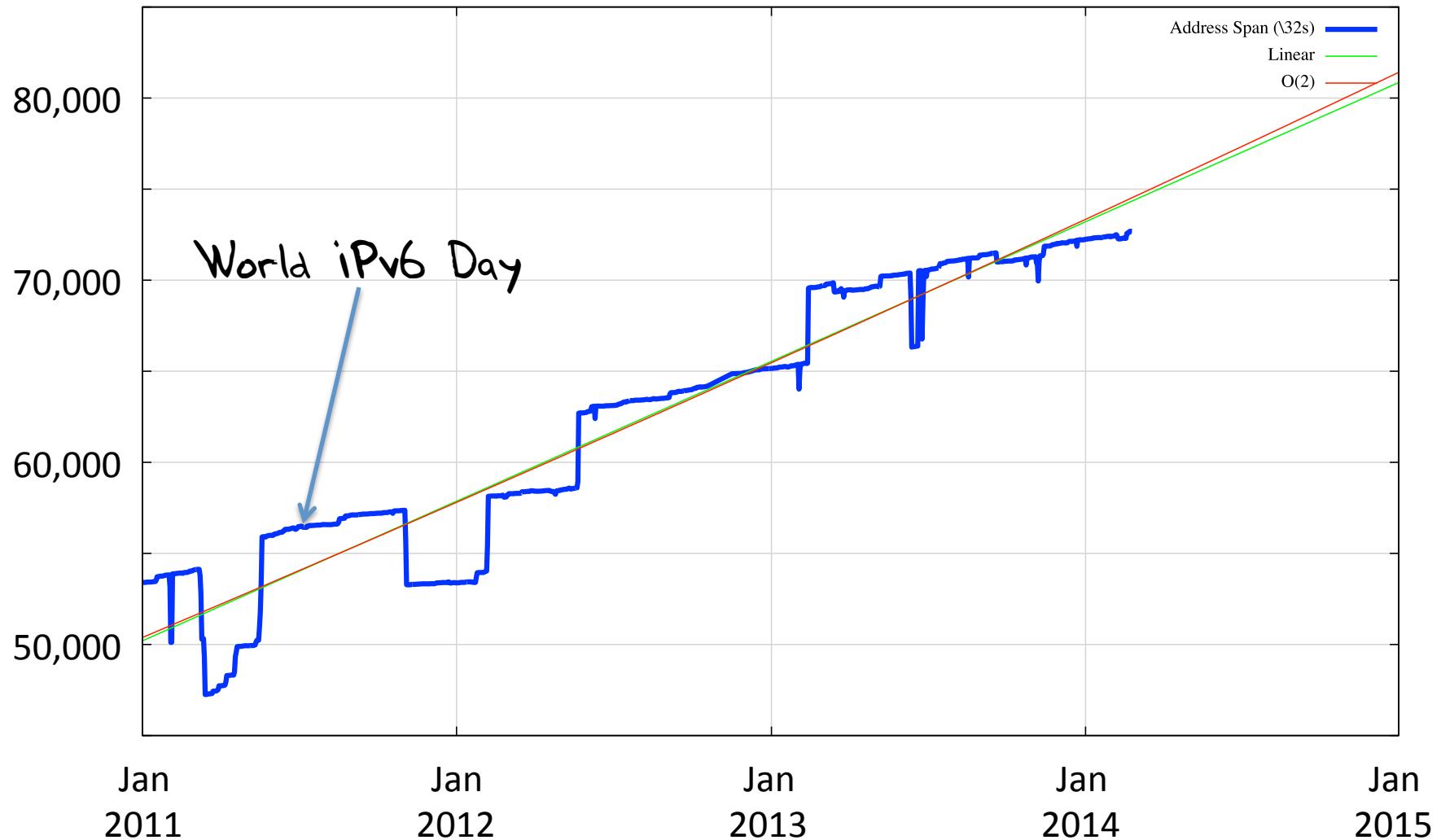
- Overall Internet growth in terms of BGP is at a rate of some **~8-10% p.a.**
- Address span growing far more slowly than the table size
- The rate of growth of the IPv4 Internet is slowing down
  - Address shortages?
  - Masking by NAT deployments?
  - Saturation of critical market sectors?

# IPv6 BGP Prefix Count

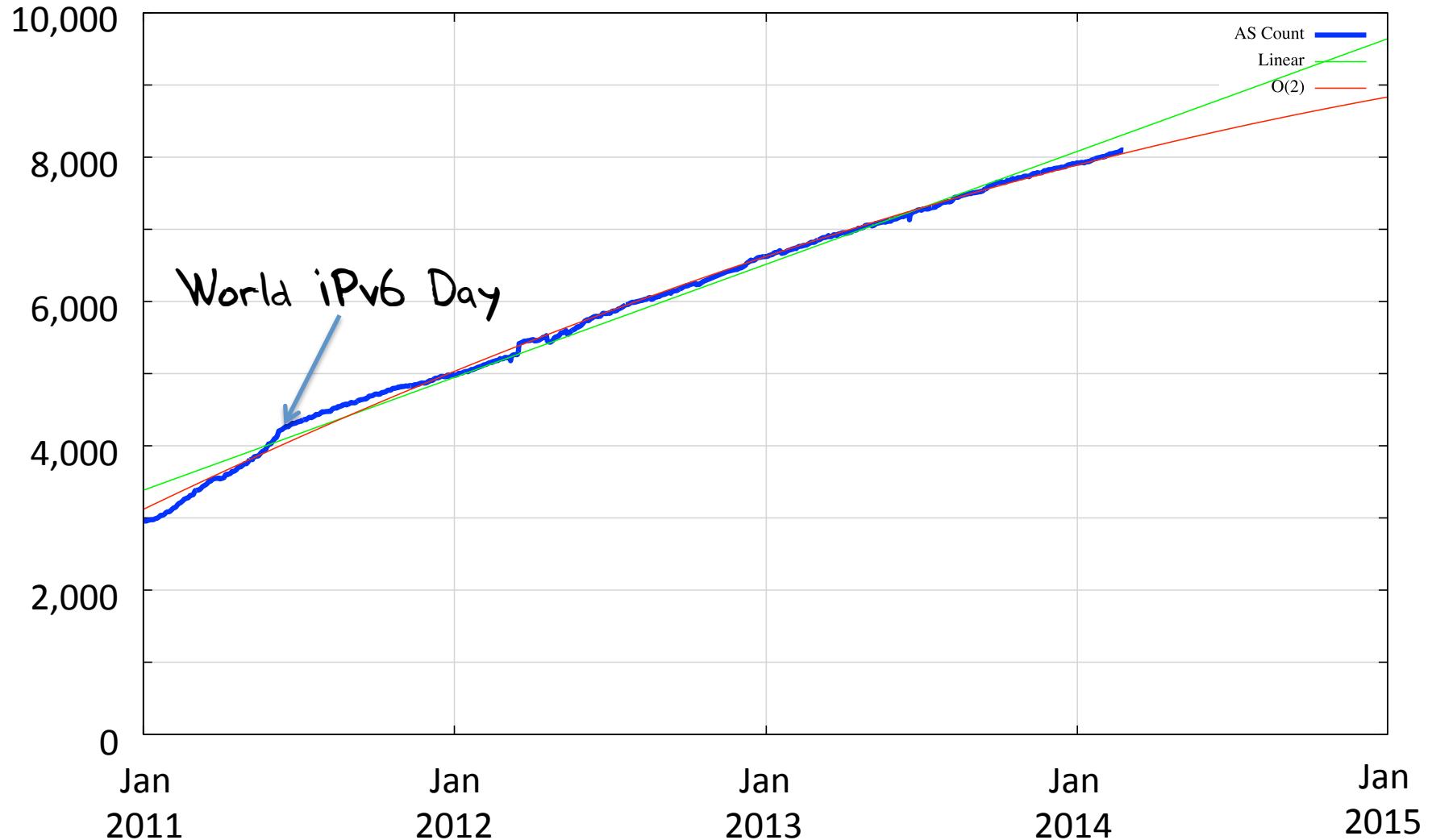
V6 BGP FIB Size



# IPv6 Routed Address Span



# IPv6 Routed AS Count



# IPv6 2011 BGP Vital Statistics

	Jan-13	Jan-14	p.a. rate
<b>Prefix Count</b>	11,500	16,100	+ 40%
Roots	8,451	11,301	+ 34%
More Specifics	3,049	4,799	+ 57%
<b>Address Span (/32s)</b>	65,127	72,245	+ 11%
<b>AS Count</b>	6,560	7,845	+ 20%
Transit	1,260	1,515	+ 20%
Stub	5,300	6,330	+ 19%

# IPv6 in 2013

- Overall IPv6 Internet growth in terms of BGP is  
**20% - 40 % p.a.**
  - 2012 growth rate was ~ 90%.

(Looking at the AS count, if these relative growth rates persist then the IPv6 network would span the same network domain as IPv4 in 16 years time -- 2030!)

# IPv6 in 2013 – Growth is Slowing

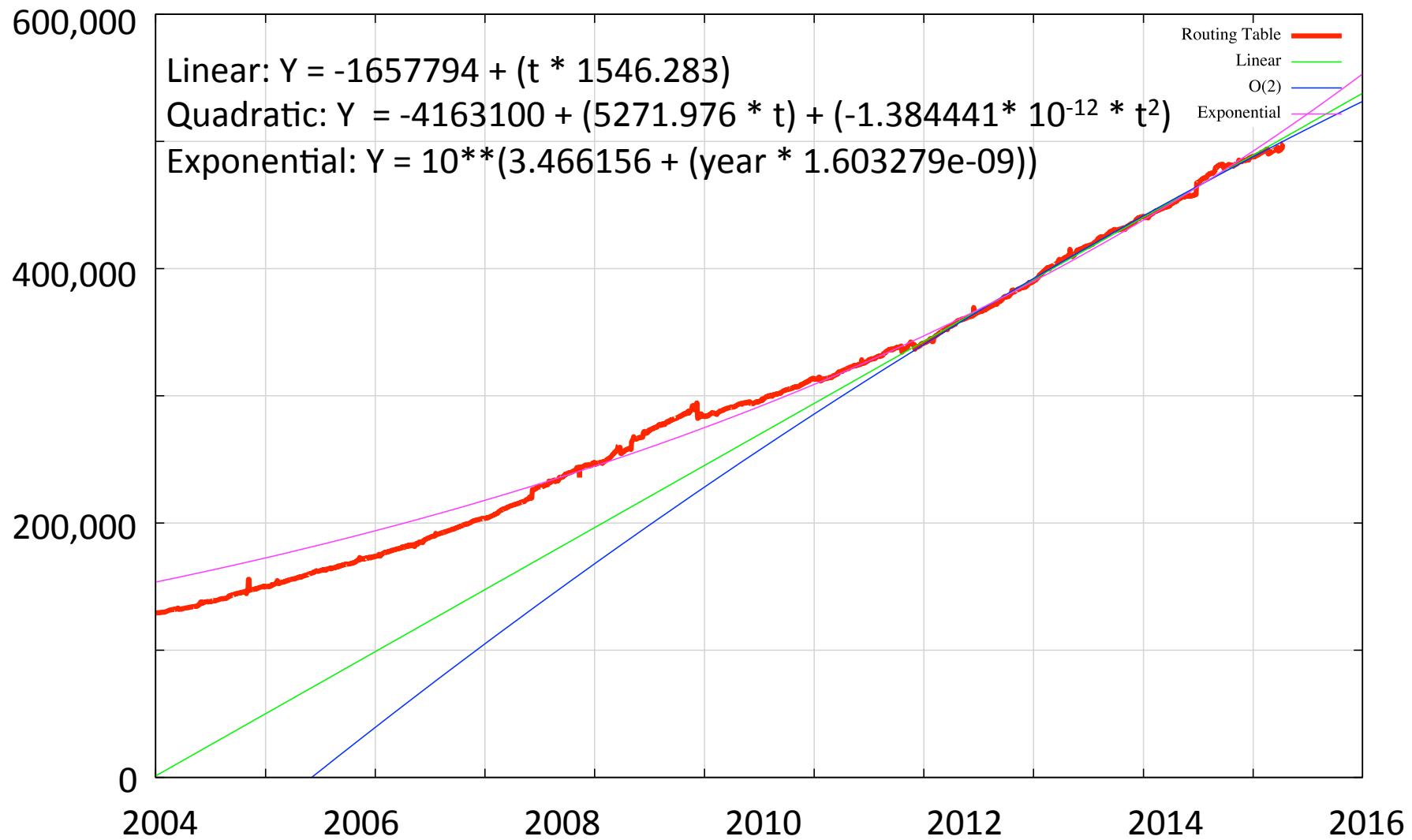
- Overall Internet growth in terms of BGP is at a rate of some **~20-40% p.a.**
- AS growth sub-linear
- The rate of growth of the IPv6 Internet is also slowing down
  - Lack of critical momentum behind IPv6?
  - Saturation of critical market sectors by IPv4?
  - *<some other factor>*?

# What to expect

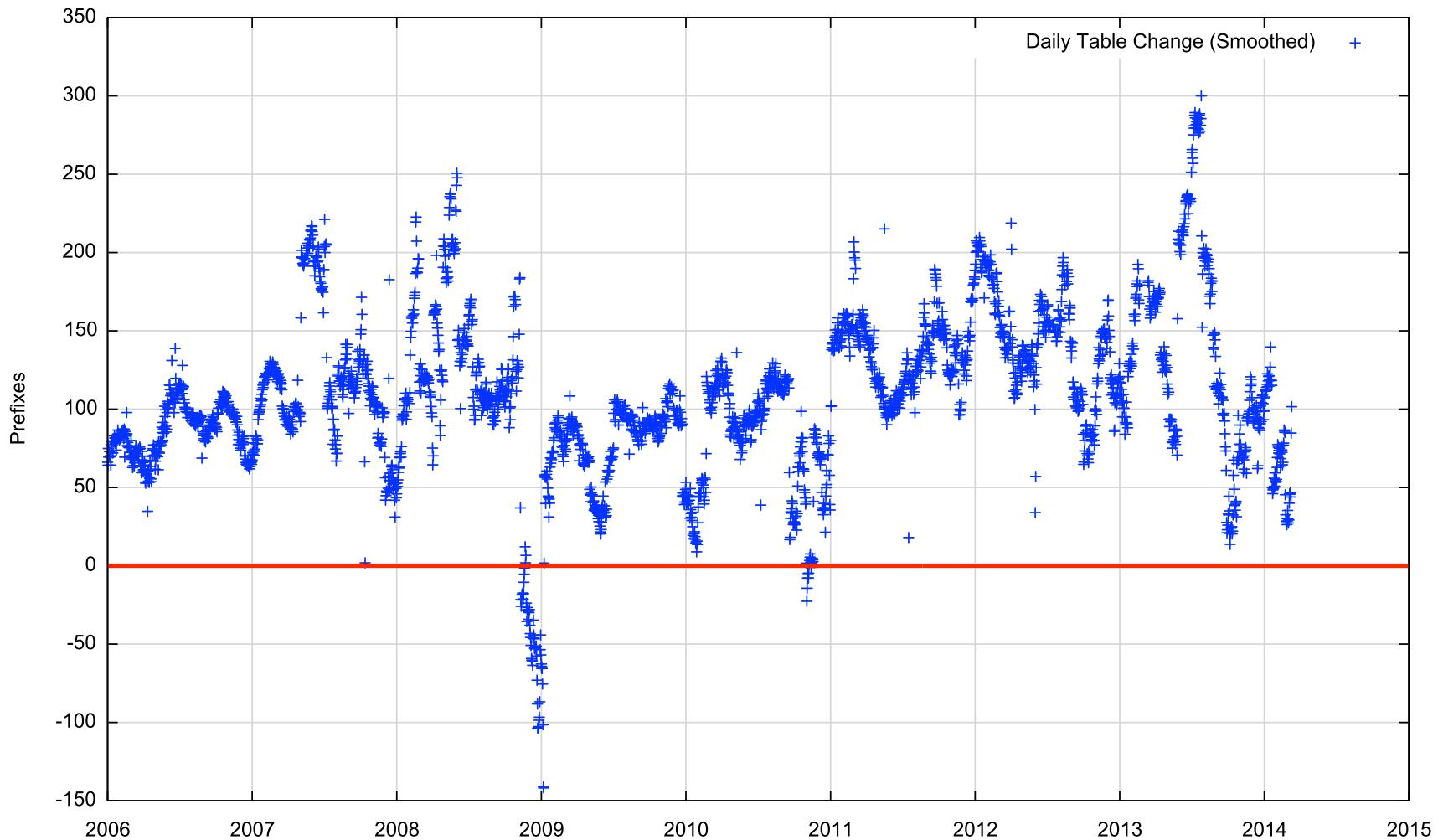
# BGP Size Projections

- Generate a projection of the IPv4 routing table using a quadratic ( $O(2)$  polynomial) over the historic data
    - For IPv4 this is a time of **extreme uncertainty**
      - Registry IPv4 address run out
      - Uncertainty over the impacts of any after-market in IPv4 on the routing table
- which makes this projection even more speculative than normal!

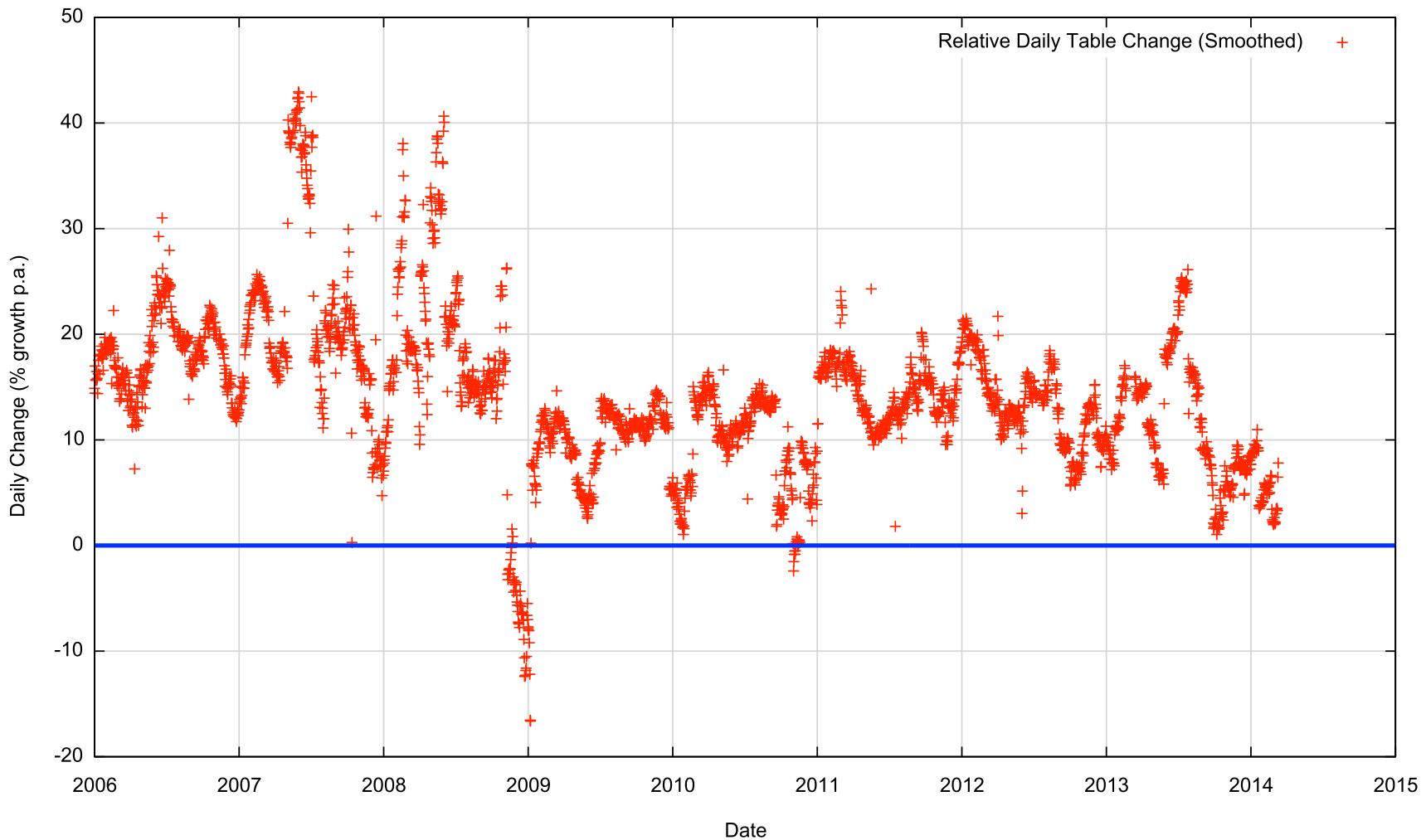
# IPv4 Table Size



# V4 - Daily Growth Rates



# V4 - Relative Growth Rates

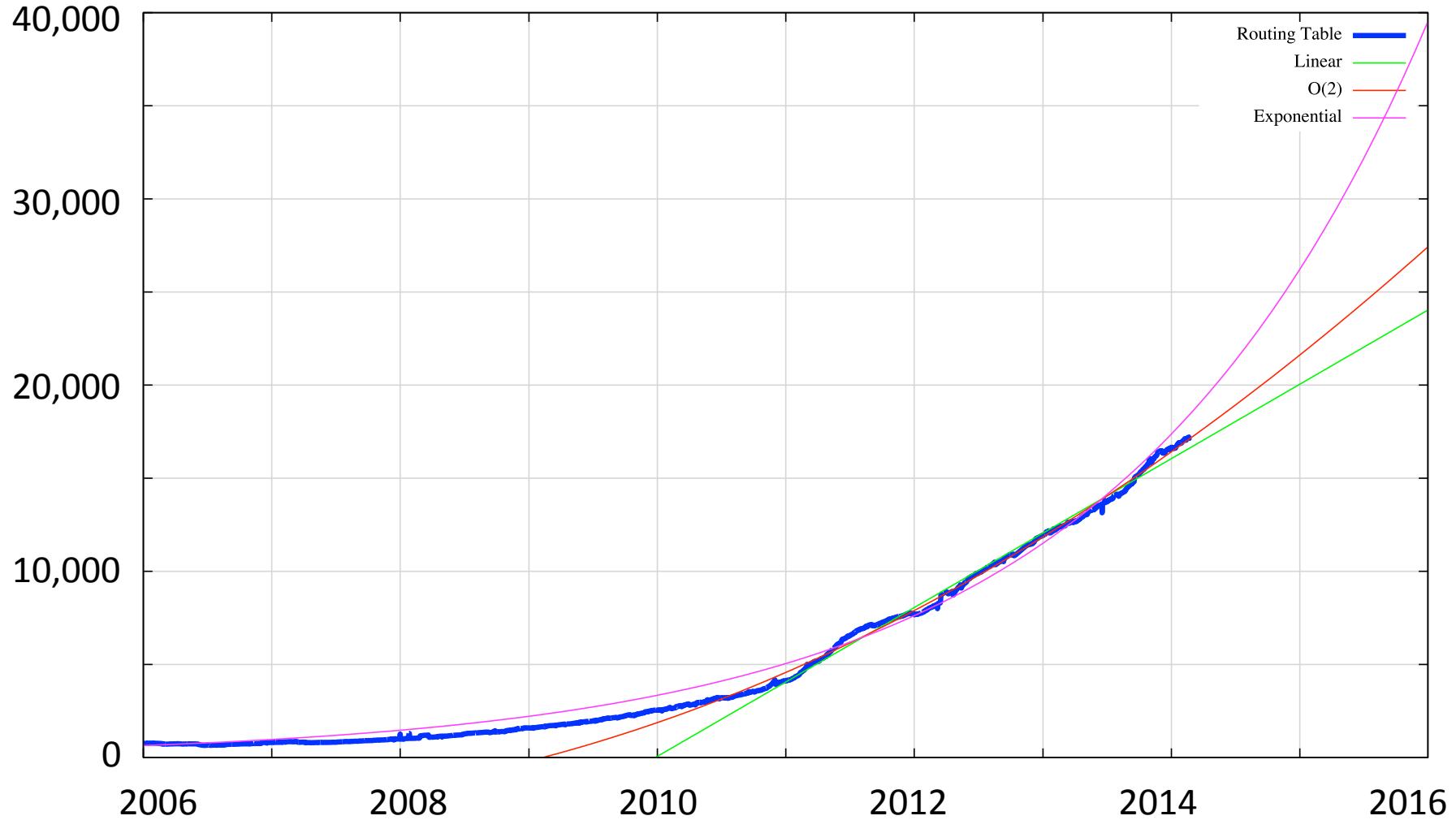


# IPv4 BGP Table Size predictions

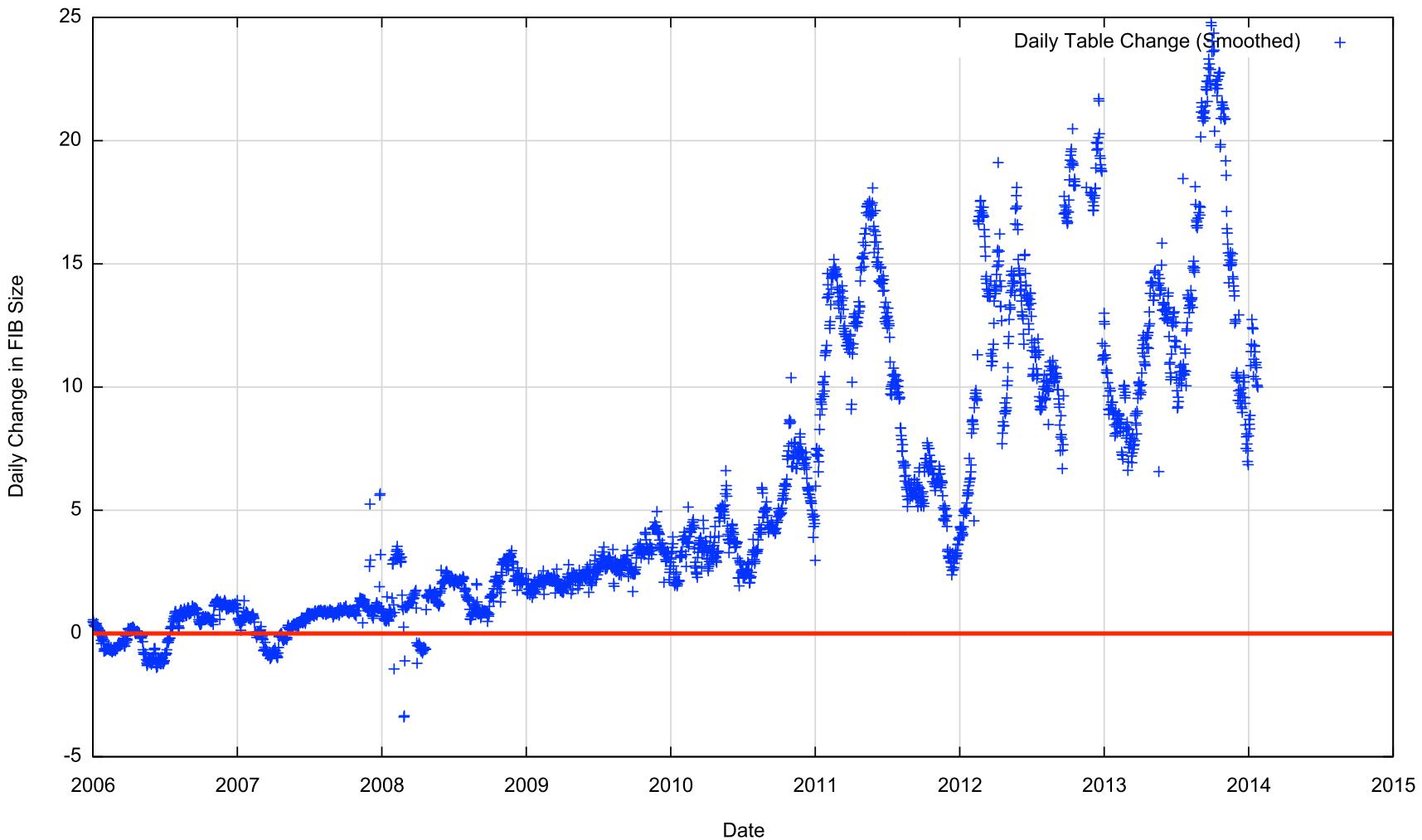
	Linear Model	Exponential Model
Jan 2013	441,172 entries	
2014	488,011 entries	
2015	540,000 entries	559,000
2016	590,000 entries	630,000
2017	640,000 entries	710,000
2018	690,000 entries	801,000
2019	740,000 entries	902,000

*These numbers are dubious due to uncertainties introduced by IPv4 address exhaustion pressures.*

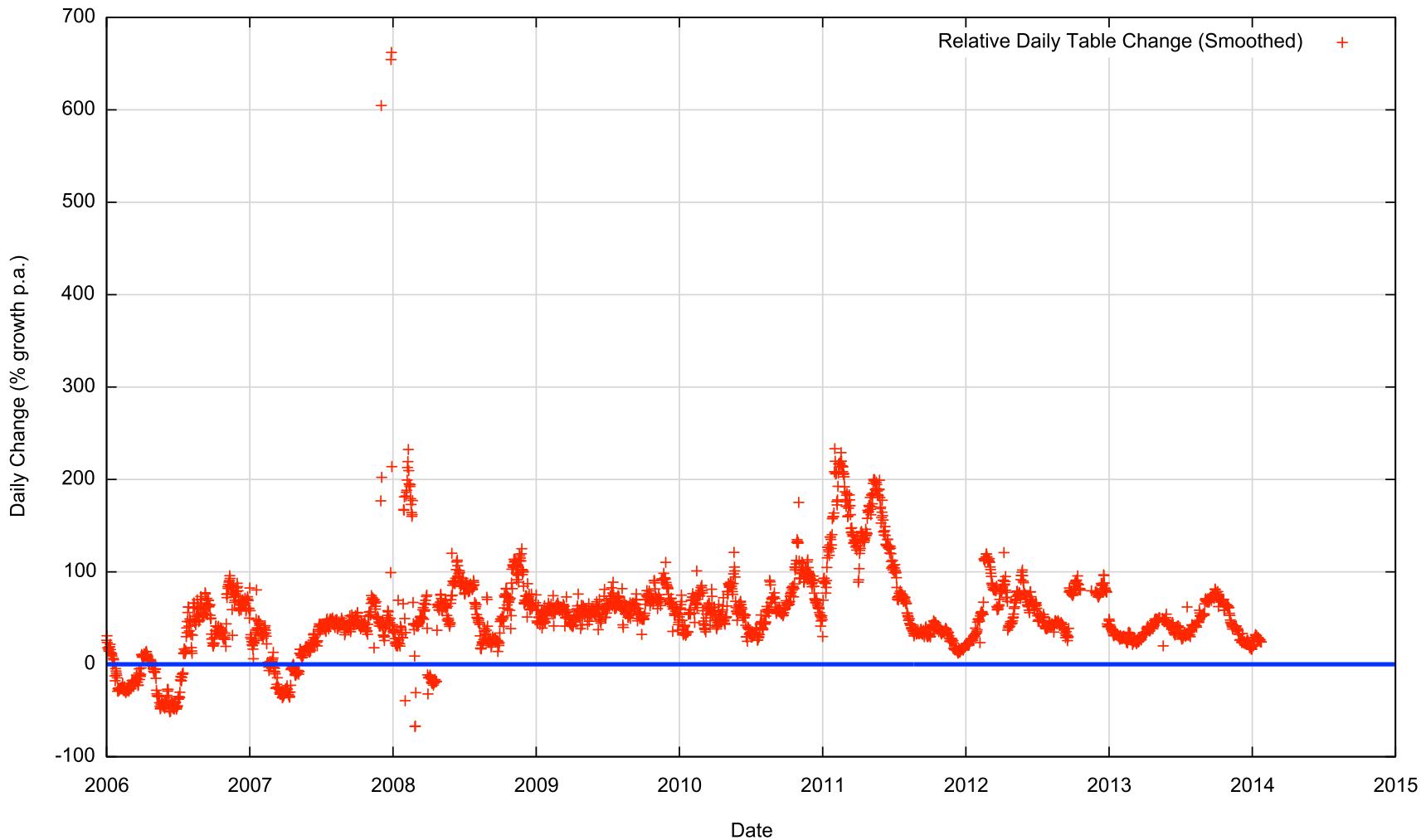
# IPv6 Table Size



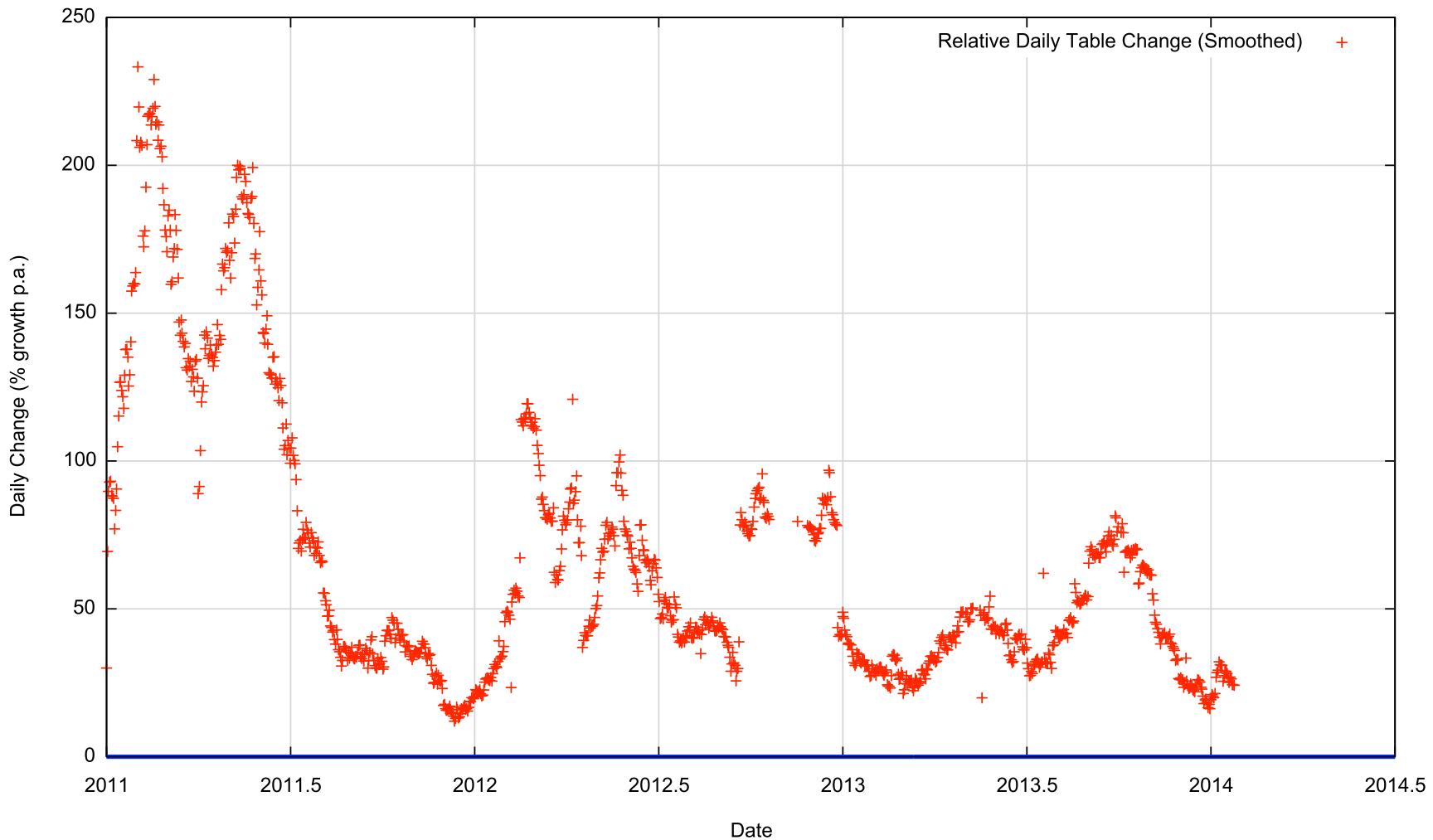
# V6 - Daily Growth Rates



# V6 - Relative Growth Rates



# V6 - Relative Growth Rates



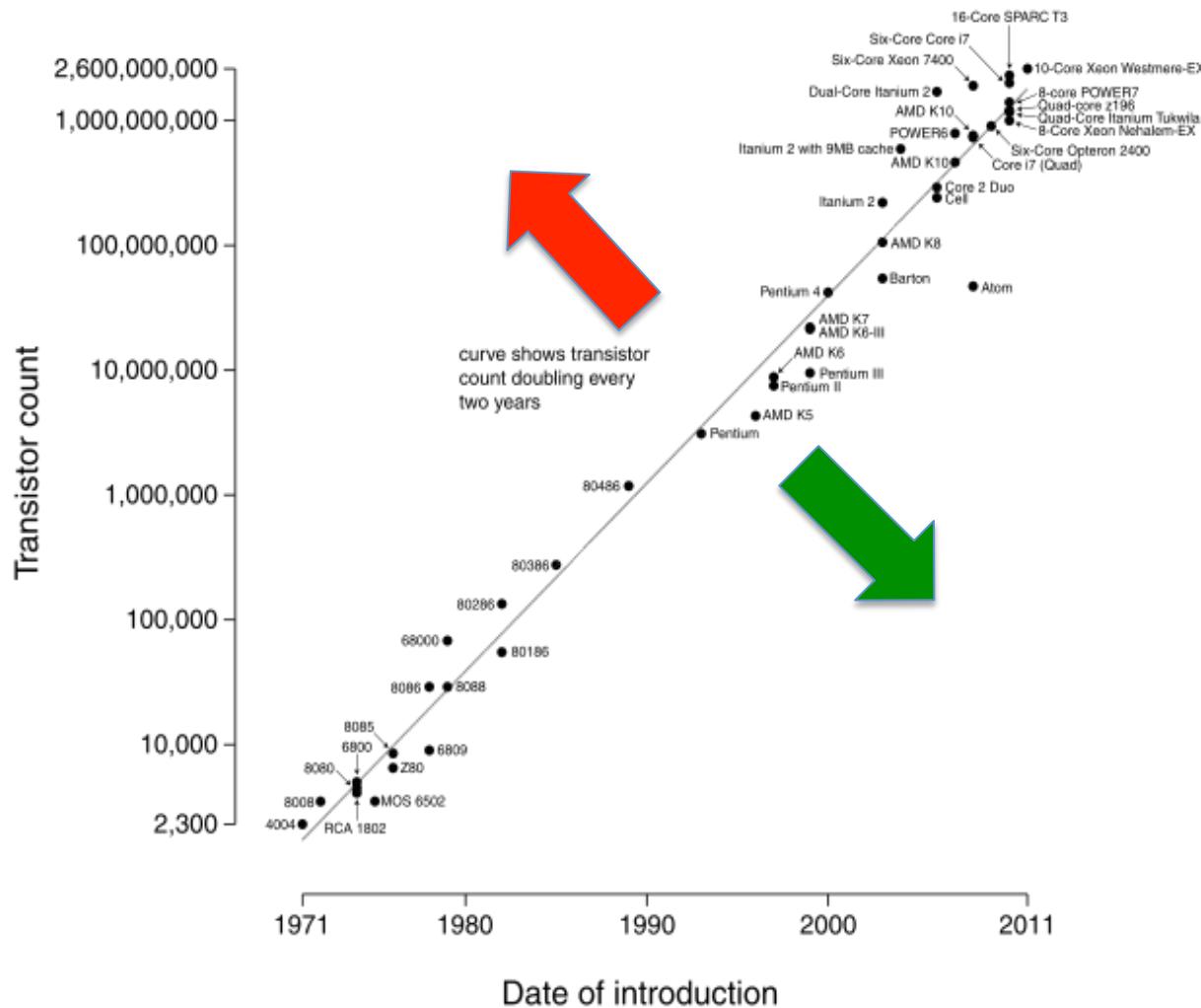
# IPv6 BGP Table Size predictions

	Exponential Model	Linear Model
Jan 2013	11,600 entries	
2014	16,200 entries	
2015	25,400 entries	19,000
2016	38,000 entries	23,000
2017	57,000 entries	27,000
2018	85,000 entries	30,000
2019	127,000 entries	35,000

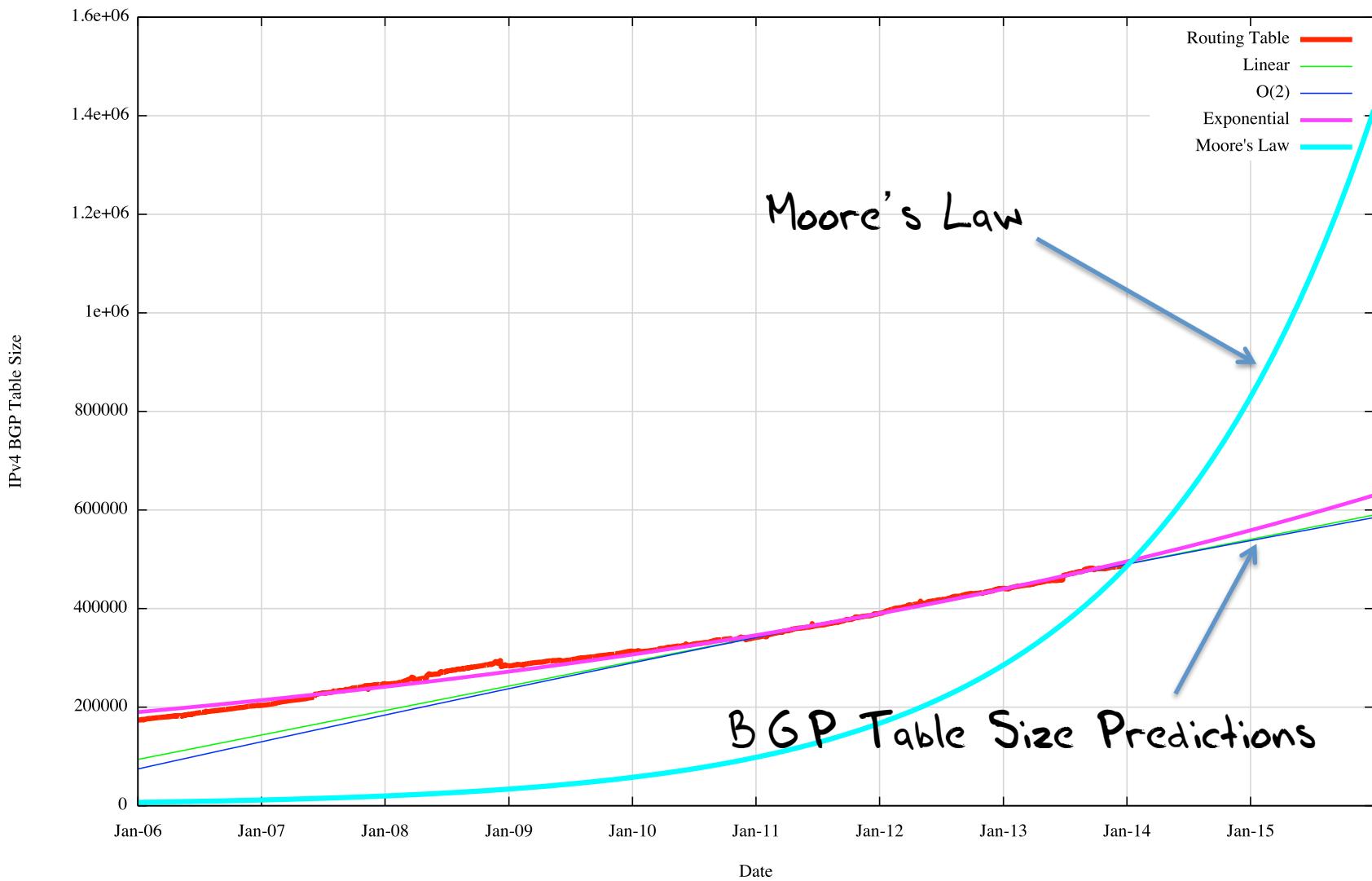
# Up and to the Right

- Most Internet curves are “up and to the right”
- But what makes this curve painful?
  - The pain threshold is approximated by Moore’s Law

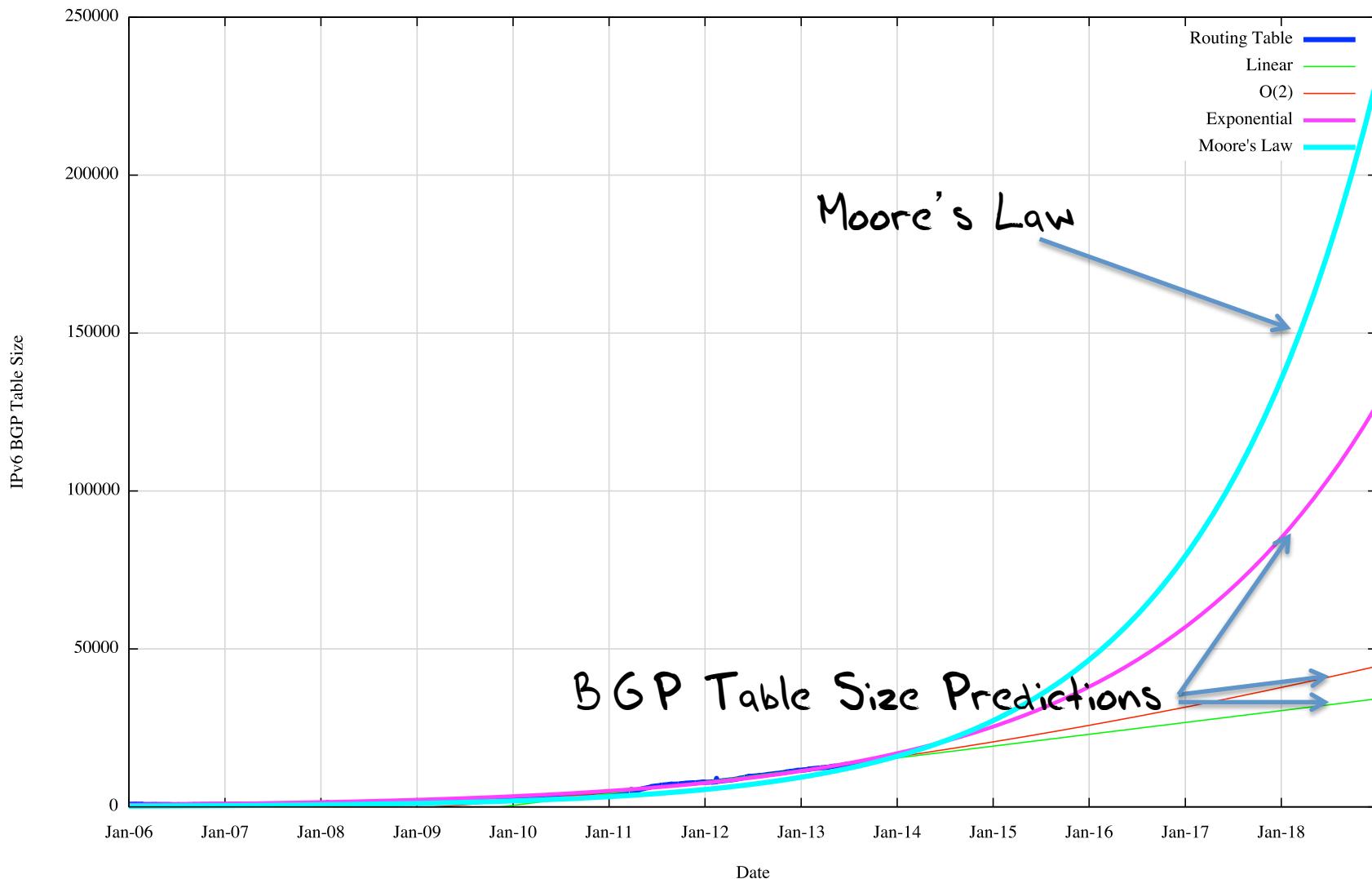
## Microprocessor Transistor Counts 1971-2011 & Moore's Law



# IPv4 BGP Table size and Moore's Law



# IPv6 Projections and Moore's Law



# eBGP Table Growth

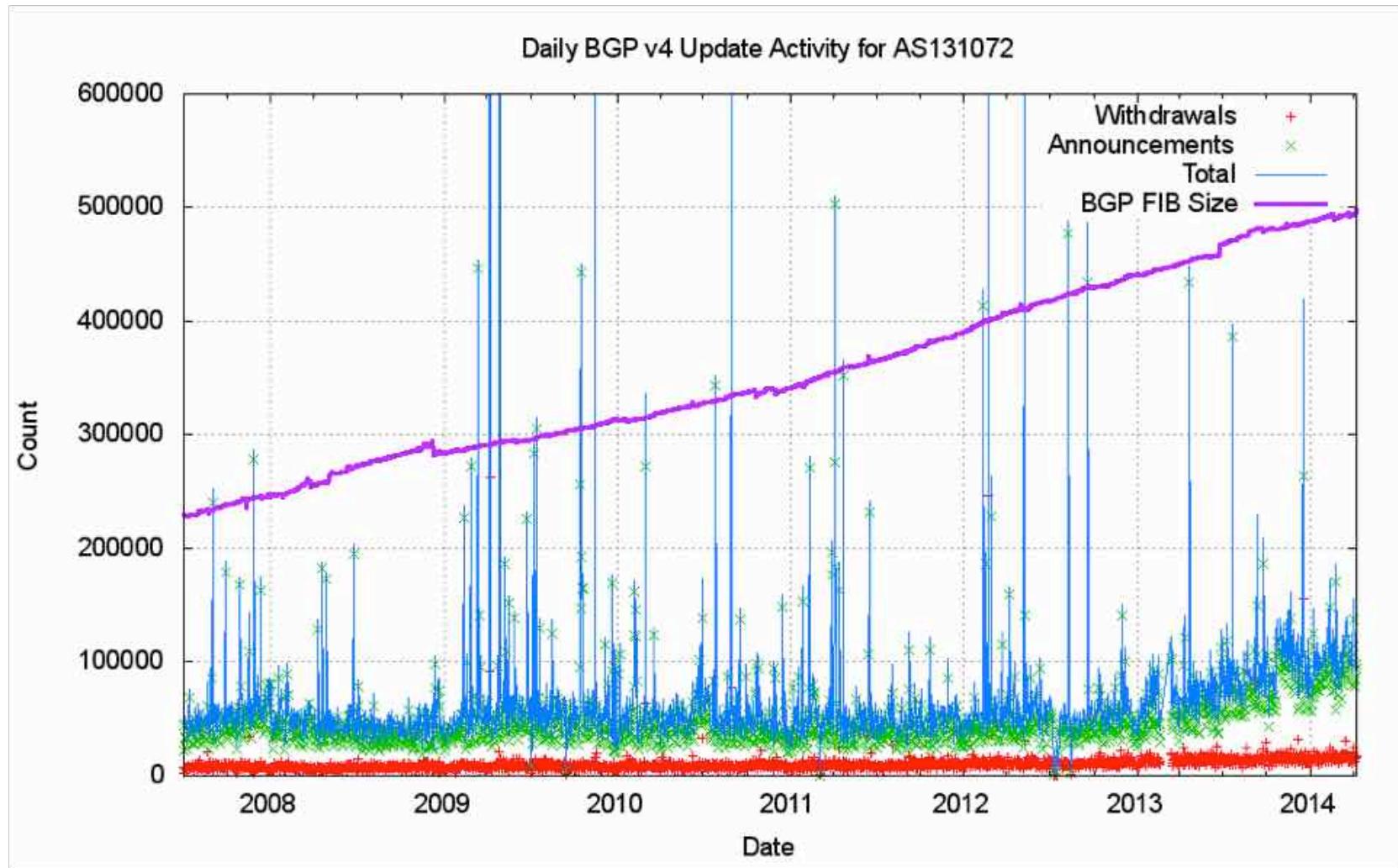
- Nothing in these figures suggests that there is cause for urgent alarm -- at present
- The overall eBGP growth rates for IPv4 are holding at a modest level, and the IPv6 table, although it is growing rapidly, is still relatively small in size in absolute terms
- As long as we are prepared to live within the technical constraints of the current routing paradigm it will continue to be viable for some time yet

# Table Size vs Updates

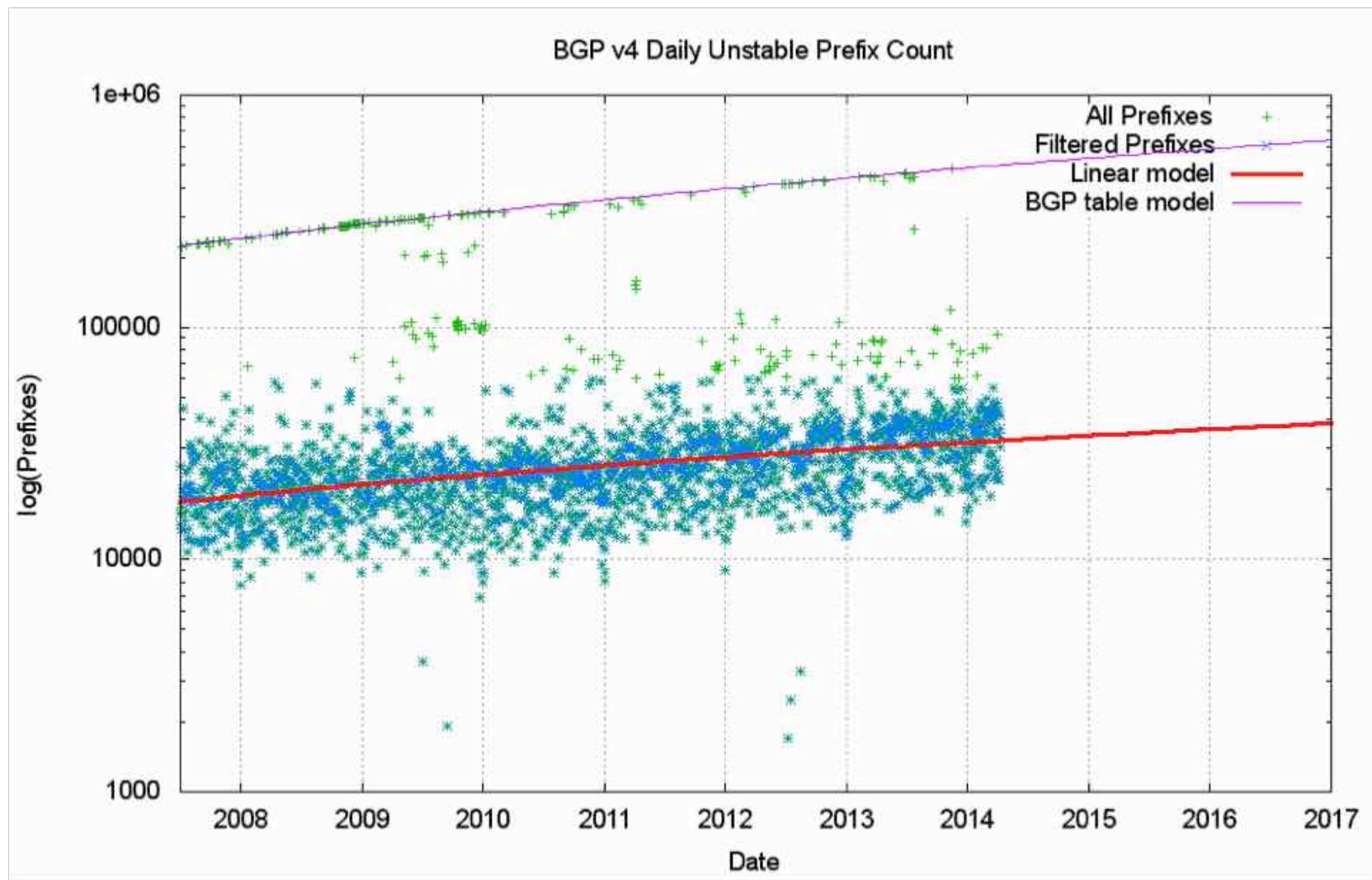
# BGP Updates

- What about the level of updates in BGP?
- Let's look at the update load from a single eBGP feed in a DFZ context

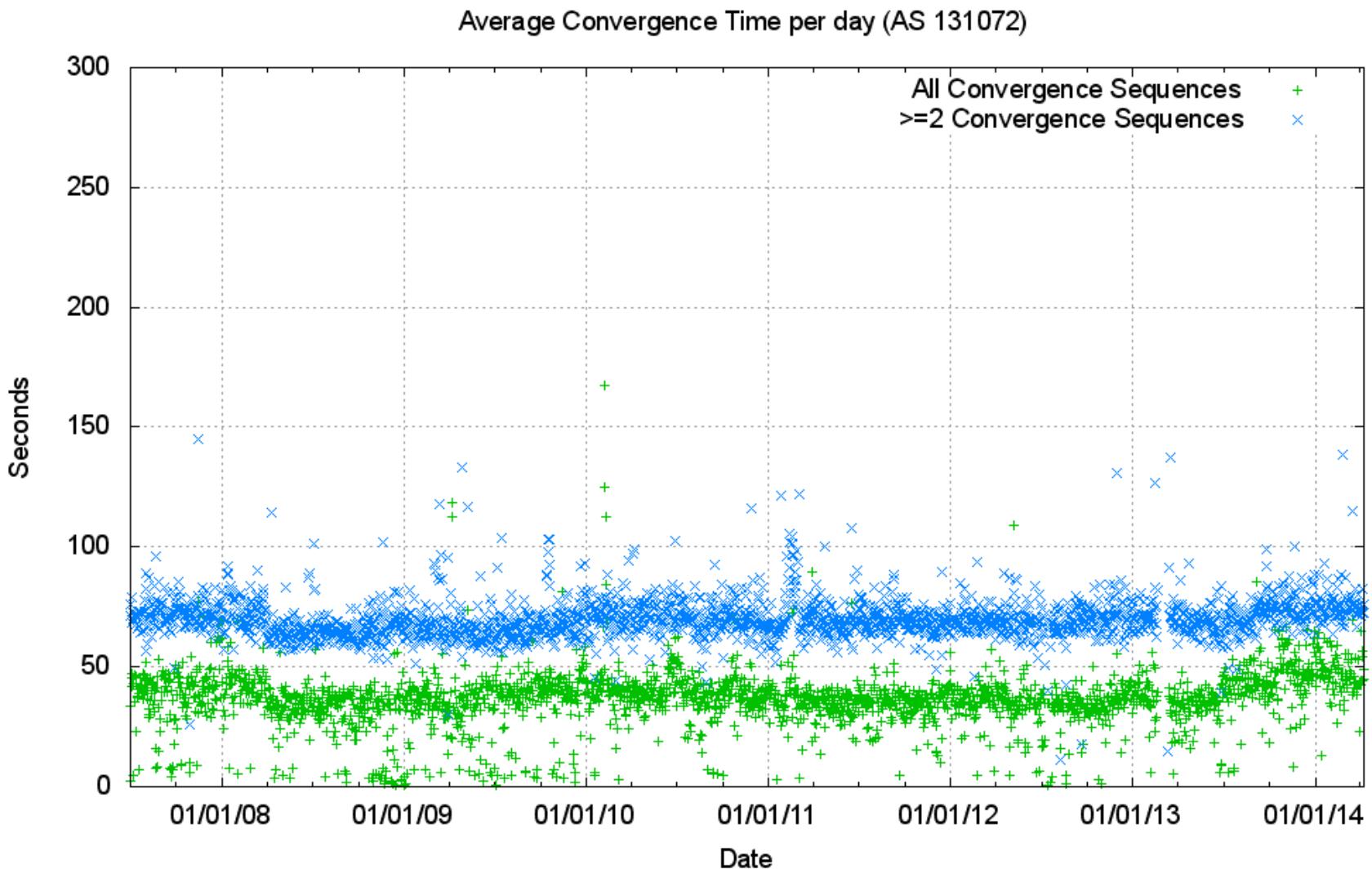
# Announcements and Withdrawals



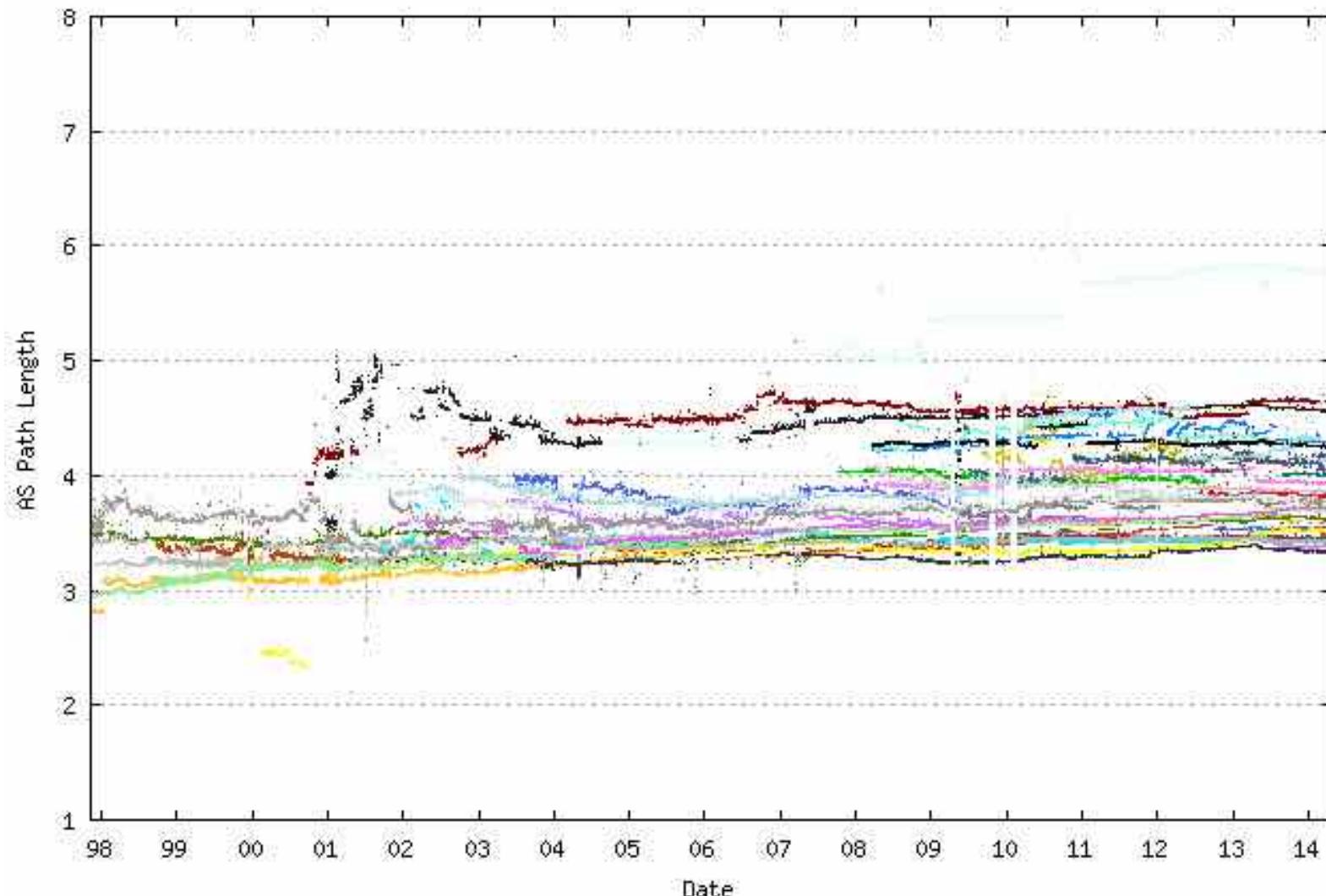
# Unstable Prefixes



# Convergence Performance



# IPv4 Average AS Path Length



Data from Route Views

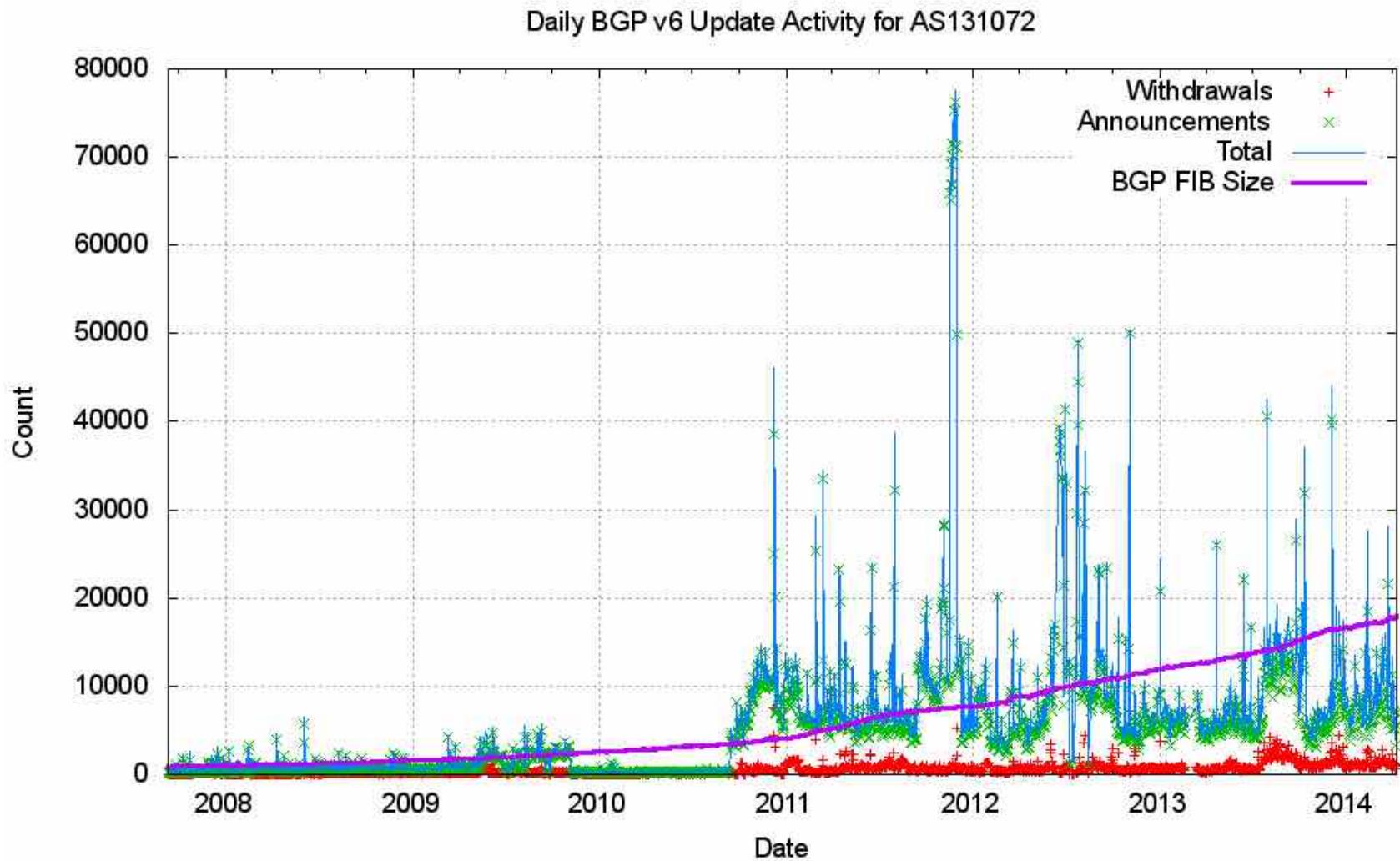
# Updates in IPv4 BGP

Nothing in these figures is cause for any great level of concern ...

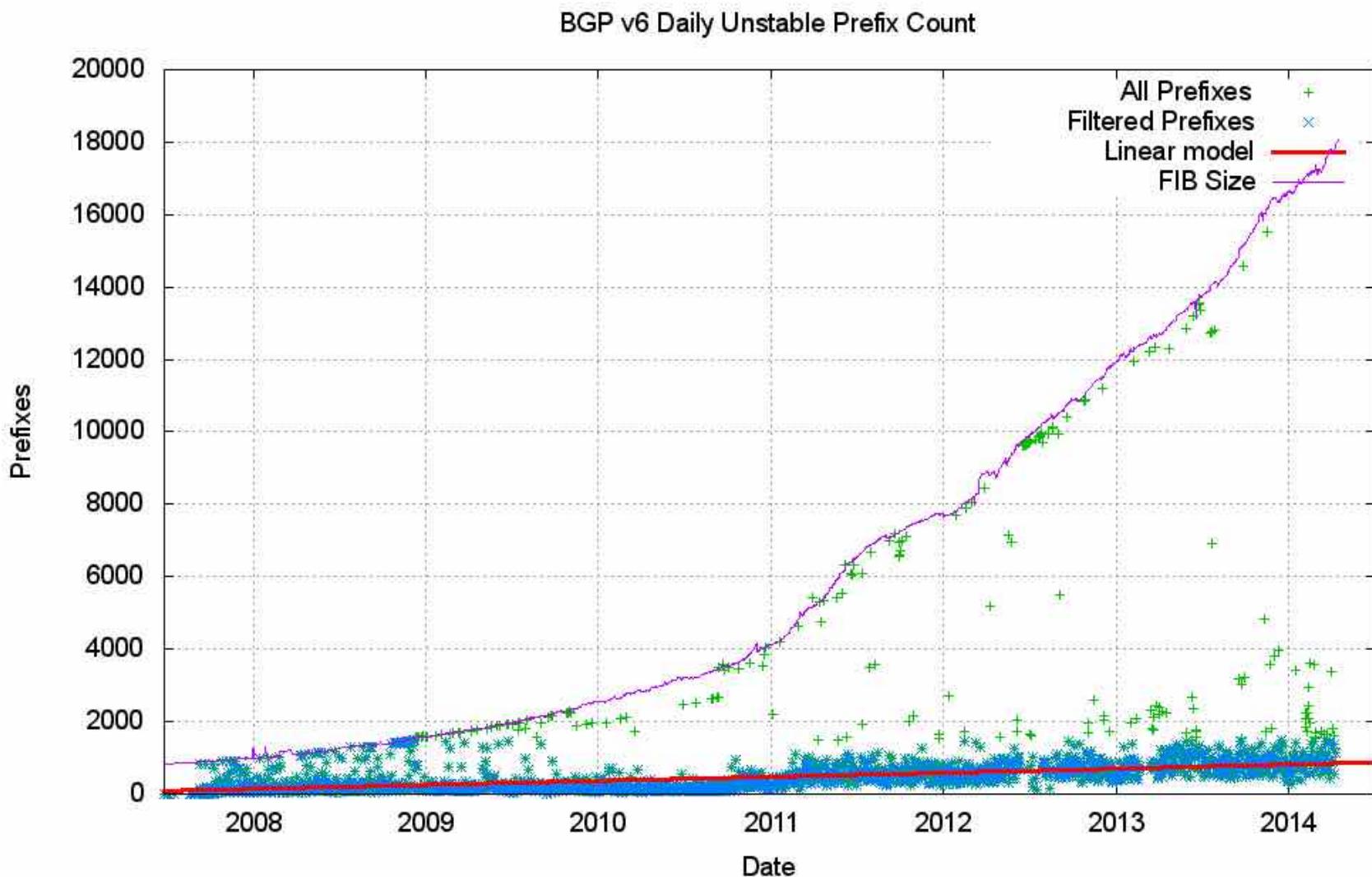
- The number of unstable prefixes per day is growing at a far lower rate than the number of announced prefixes
- The number of updates per instability event has been constant, due to the damping effect of the MRAI interval, and the relatively constant AS Path length over this interval

What about IPv6?

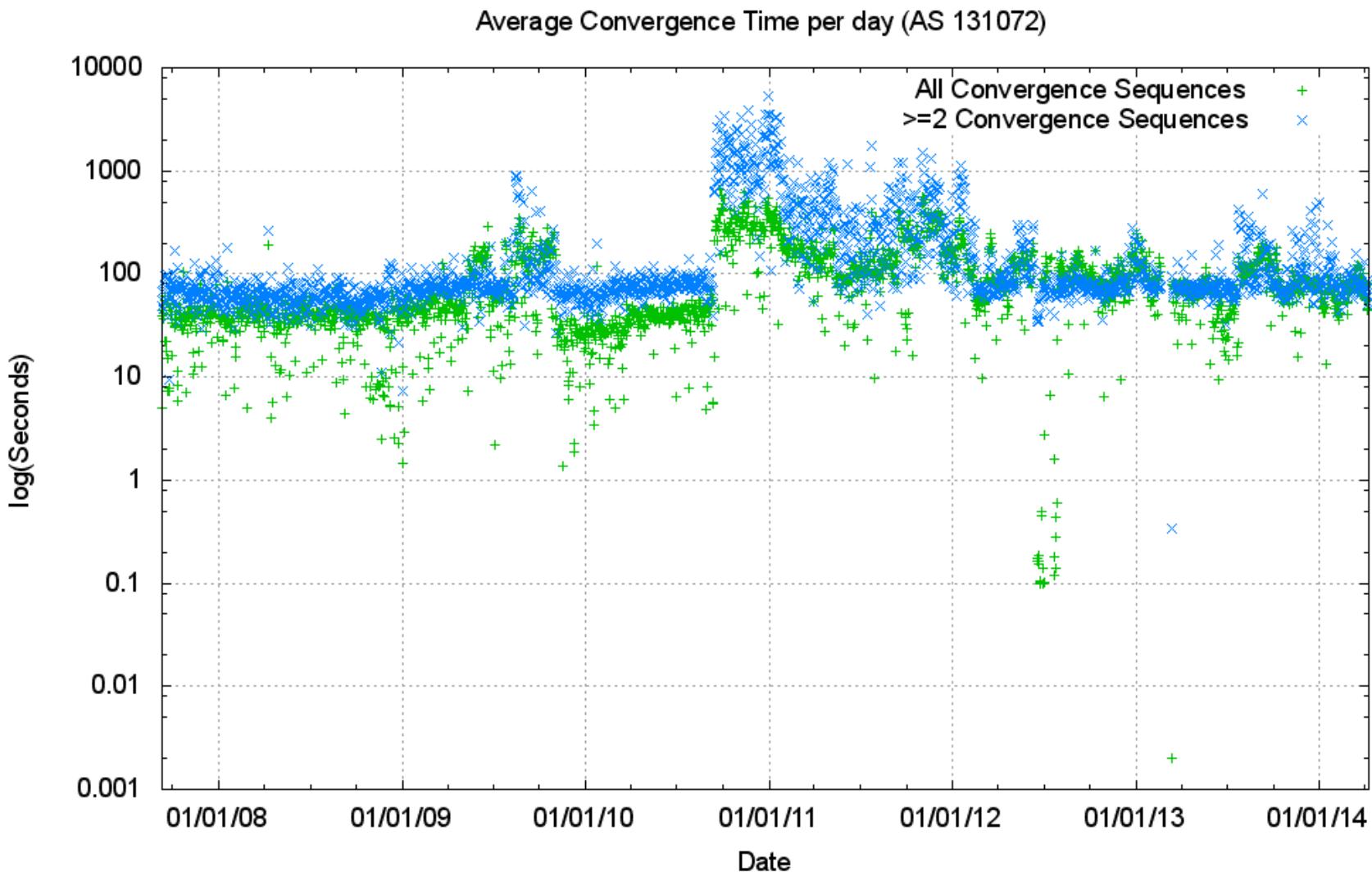
# V6 Announcements and Withdrawals



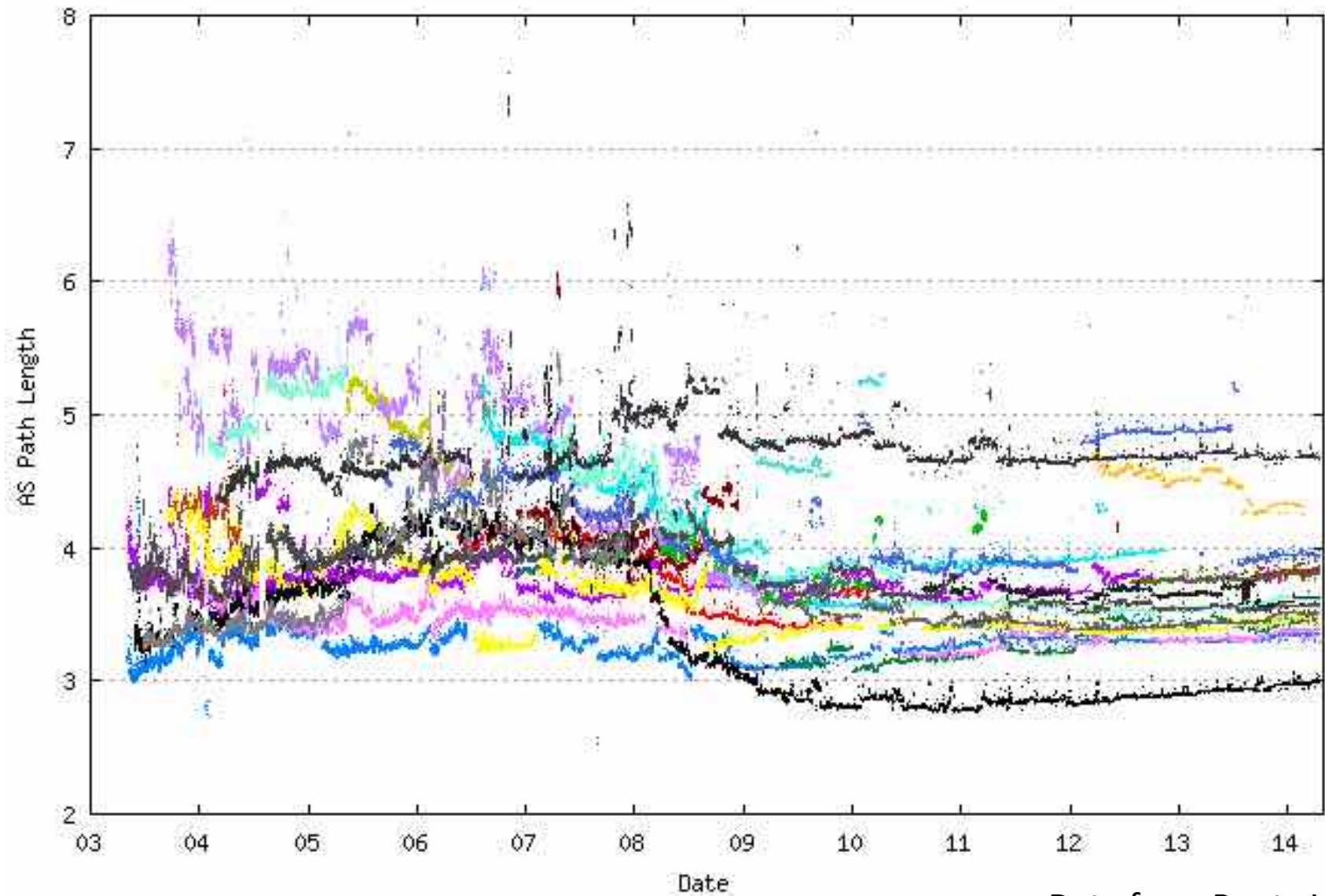
# Unstable Prefixes



# Convergence Performance



# IPv6 Average AS Path Length



# BGP Convergence

- The long term average convergence time for the IPv4 BGP network is some 70 seconds, or 2.3 updates given a 30 second MRAI timer
- The long term average convergence time for the IPv6 BGP network is some 80 seconds, or 2.6 updates

# BGP Table Growth

However ... continued scalability of the routing system relies on continued conservatism in routing practices.

How good are we at “being conservative” in routing?

# CIDR and BGP

- To what extent do we still practice “conservative” routing and refrain from announcing more specifics into the routing table?
- Are we getting better or worse at aggregation in routing?
- What is the distribution of advertising more specifics? Are we seeing a significant increase in the number of more specific /24s in the routing table?

# An Example:

Prefix	AS Path							
193.124.0.0/15	4608	1221	4637	3356	20485	2118	?	
193.124.0.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.1.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.2.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.3.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.4.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.5.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.6.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.7.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.8.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.9.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.10.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.11.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.12.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.13.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.14.0/24	4608	1221	4637	3356	20485	2118	?	
193.124.15.0/24	4608	1221	4637	3356	20485	2118	?	

Origin AS: AS 2118 RELCOM-AS OOO "NPO Relcom"

# Who is doing this the most?

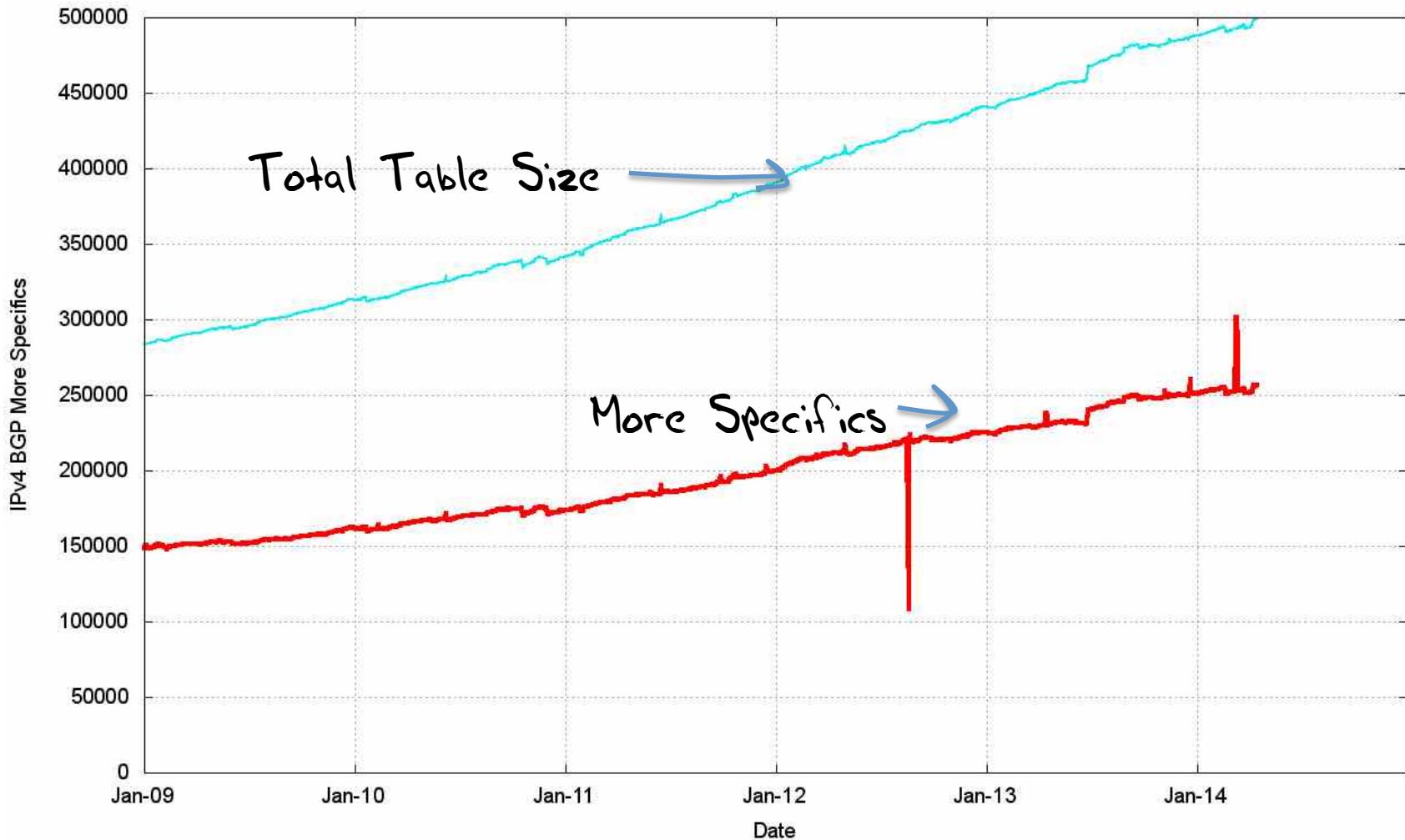
[www.cidr-report.org](http://www.cidr-report.org)

--- 16Apr14 ---

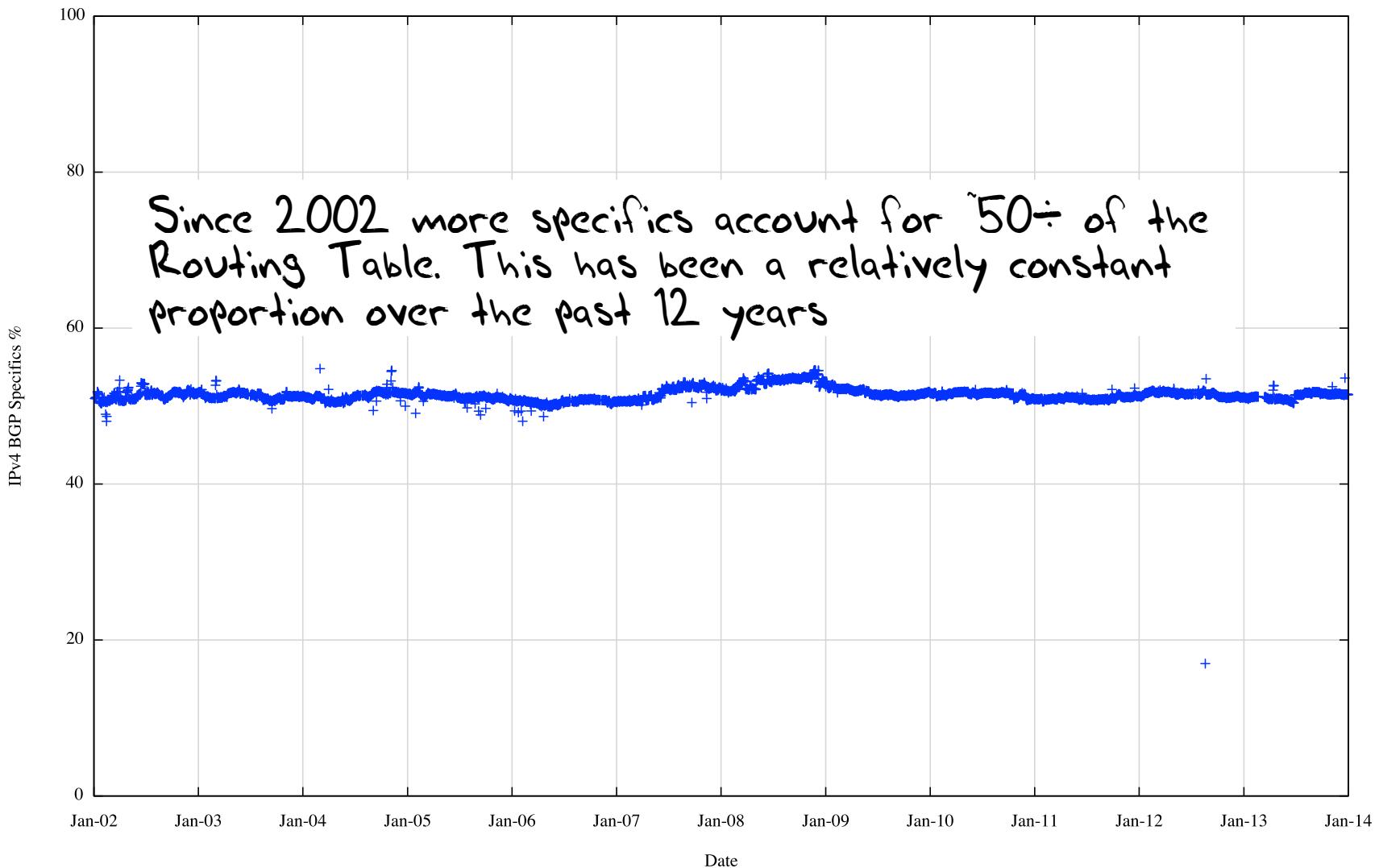
ASnum NetsNow NetsAggr NetGain % Gain Description

Table	ASnum	NetsNow	NetsAggr	NetGain	% Gain	Description
	499313	282273	217040	43.5%	All ASes	
AS28573	3701	343	3358	90.7%	NET Serviços de Comunicação S.A.,BR	
AS6389	2996	56	2940	98.1%	BELLSOUTH-NET-BLK - BellSouth.net Inc.,US	
AS17974	2781	240	2541	91.4%	TELKOMNET-AS2-AP PT Telekomunikasi Indonesia,ID	
AS4766	2919	907	2012	68.9%	KIXS-AS-KR Korea Telecom,KR	
AS18881	1924	37	1887	98.1%	Global Village Telecom,BR	
AS1785	2191	479	1712	78.1%	AS-PAETEC-NET - PaeTec Communications, Inc.,US	
AS10620	2821	1339	1482	52.5%	Telmex Colombia S.A.,CO	
AS18566	2047	565	1482	72.4%	MEGAPATH5-US - MegaPath Corporation,US	
AS36998	1637	159	1478	90.3%	SDN-MOBITEL,SD	
AS4323	2931	1511	1420	48.4%	TWTC - tw telecom holdings, inc.,US	
AS7303	1758	457	1301	74.0%	Telecom Argentina S.A.,AR	
AS4755	1843	608	1235	67.0%	TATACOMM-AS TATA Communications formerly VSNL is Leading ISP,IN	
AS7545	2232	1074	1158	51.9%	TPG-INTERNET-AP TPG Telecom Limited,AU	
AS7552	1224	113	1111	90.8%	VIETTEL-AS-AP Viettel Corporation,VN	
AS22561	1310	241	1069	81.6%	AS22561 - CenturyTel Internet Holdings, Inc.,US	
AS6983	1228	217	1011	82.3%	ITCDELTA - Earthlink, Inc.,US	
AS22773	2412	1474	938	38.9%	ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc.,US	
AS9829	1622	726	896	55.2%	BSNL-NIB National Internet Backbone,IN	
AS24560	1123	297	826	73.6%	AIRTELBROADBAND-AS-AP Bharti Airtel Ltd., Telemedia Services,IN	
AS4808	1205	387	818	67.9%	CHINA169-BJ CNCROUP IP network China169 Beijing Province Network,CN	
AS4788	1029	253	776	75.4%	TMNET-AS-AP TM Net, Internet Service Provider,MY	
AS18101	946	188	758	80.1%	RELIANCE-COMMUNICATIONS-IN Reliance Communications Ltd.DAKC MUMBAI,IN	
AS7738	914	184	730	79.9%	Telemar Norte Leste S.A.,BR	
AS701	1480	751	729	49.3%	UUNET - MCI Communications Services, Inc. d/b/a Verizon Business,US	
AS8151	1408	680	728	51.7%	Uninet S.A. de C.V.,MX	
AS26615	819	109	710	86.7%	Tim Celular S.A.,BR	

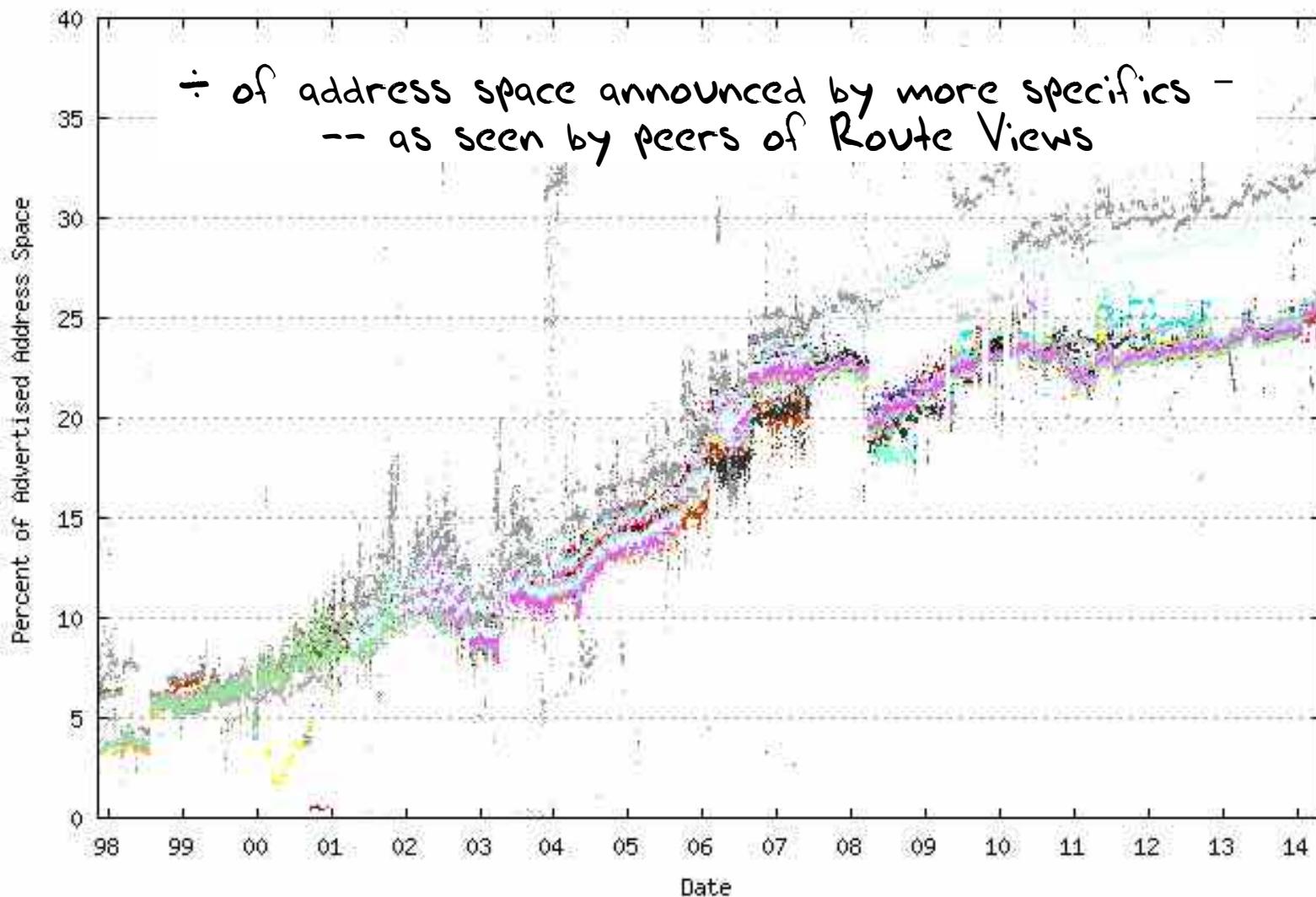
# More specifics in the Routing Table



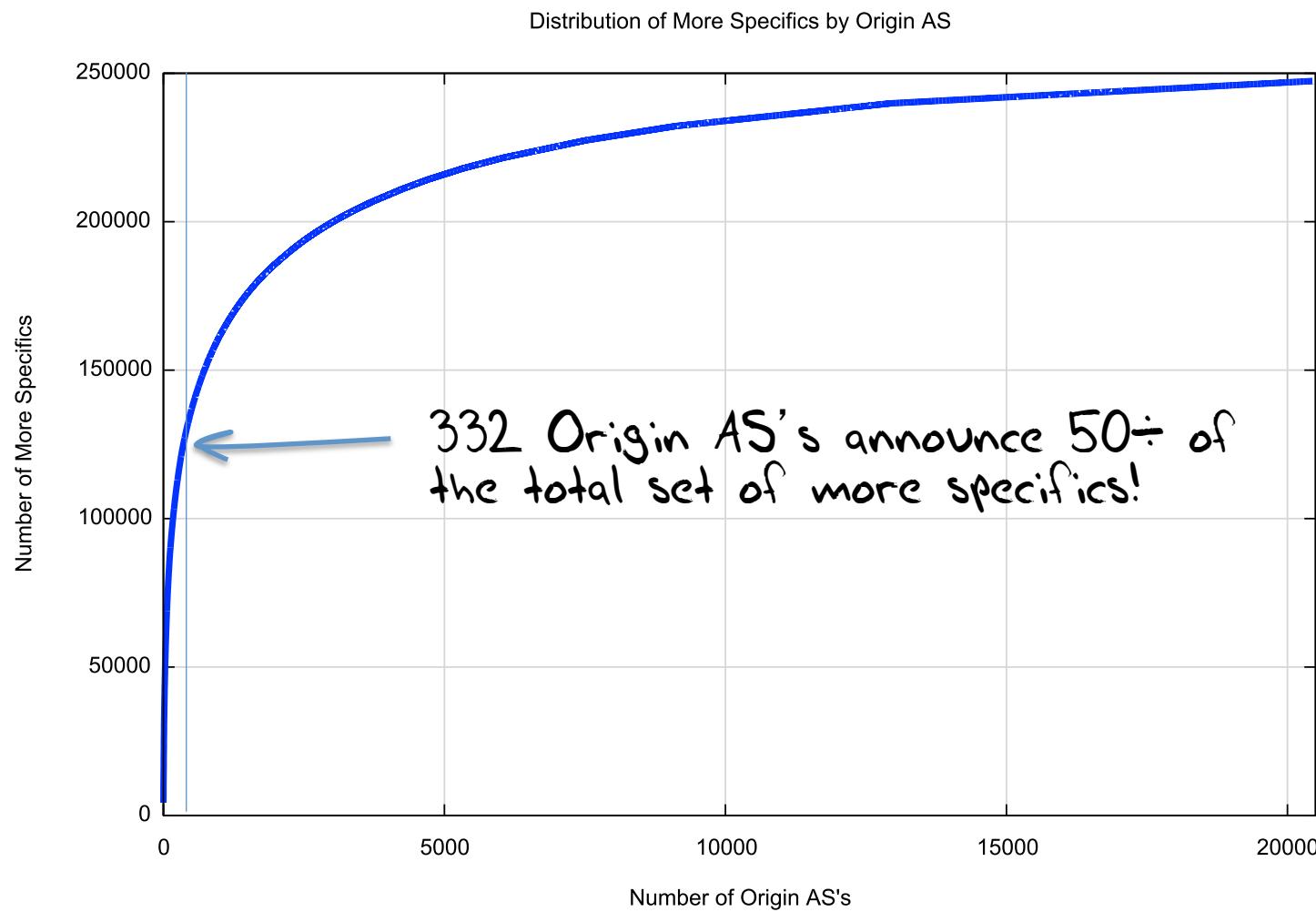
# More specifics in the Routing Table



# How much address space is announced by more specifics?



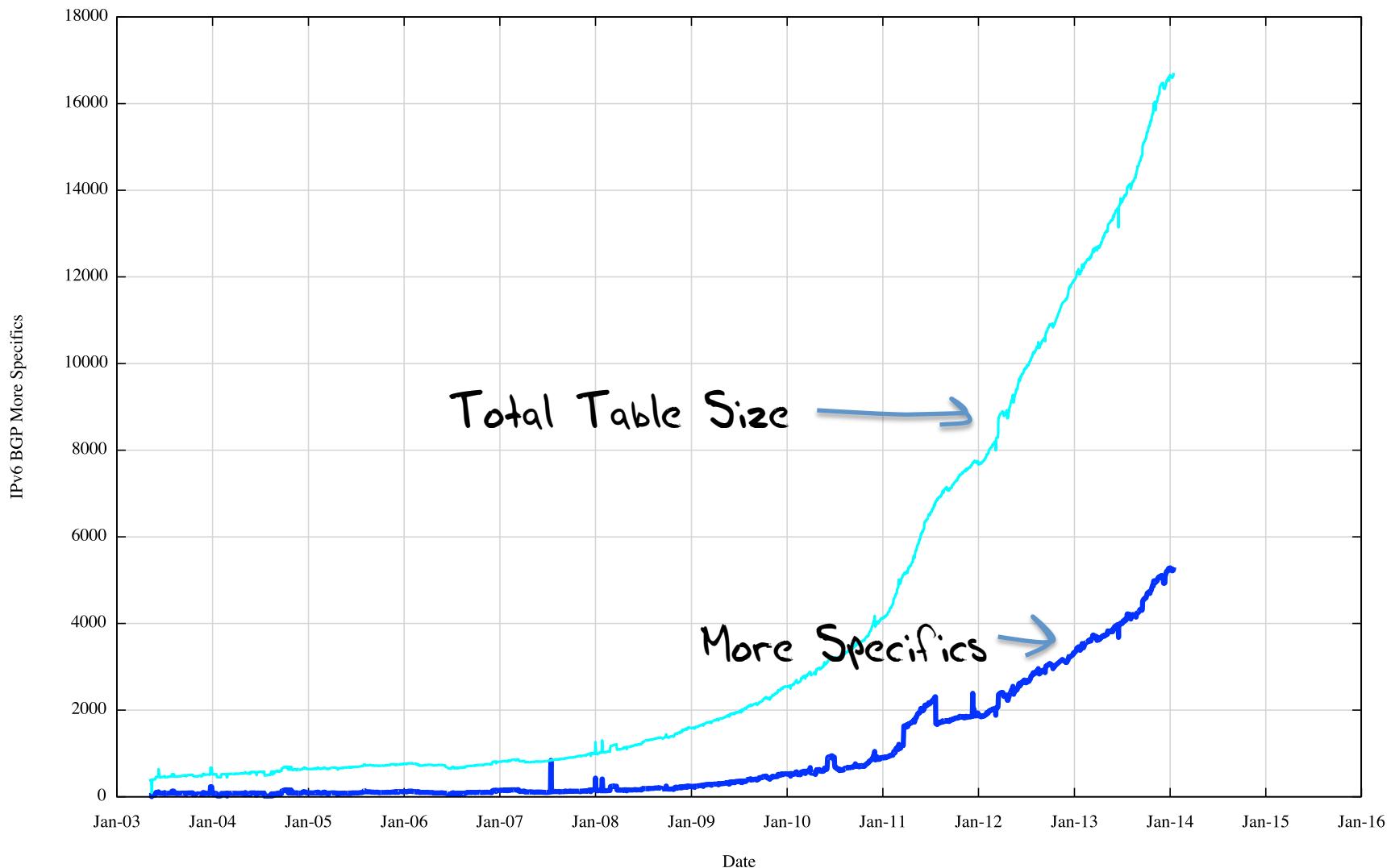
# Does everyone announce more specifics?



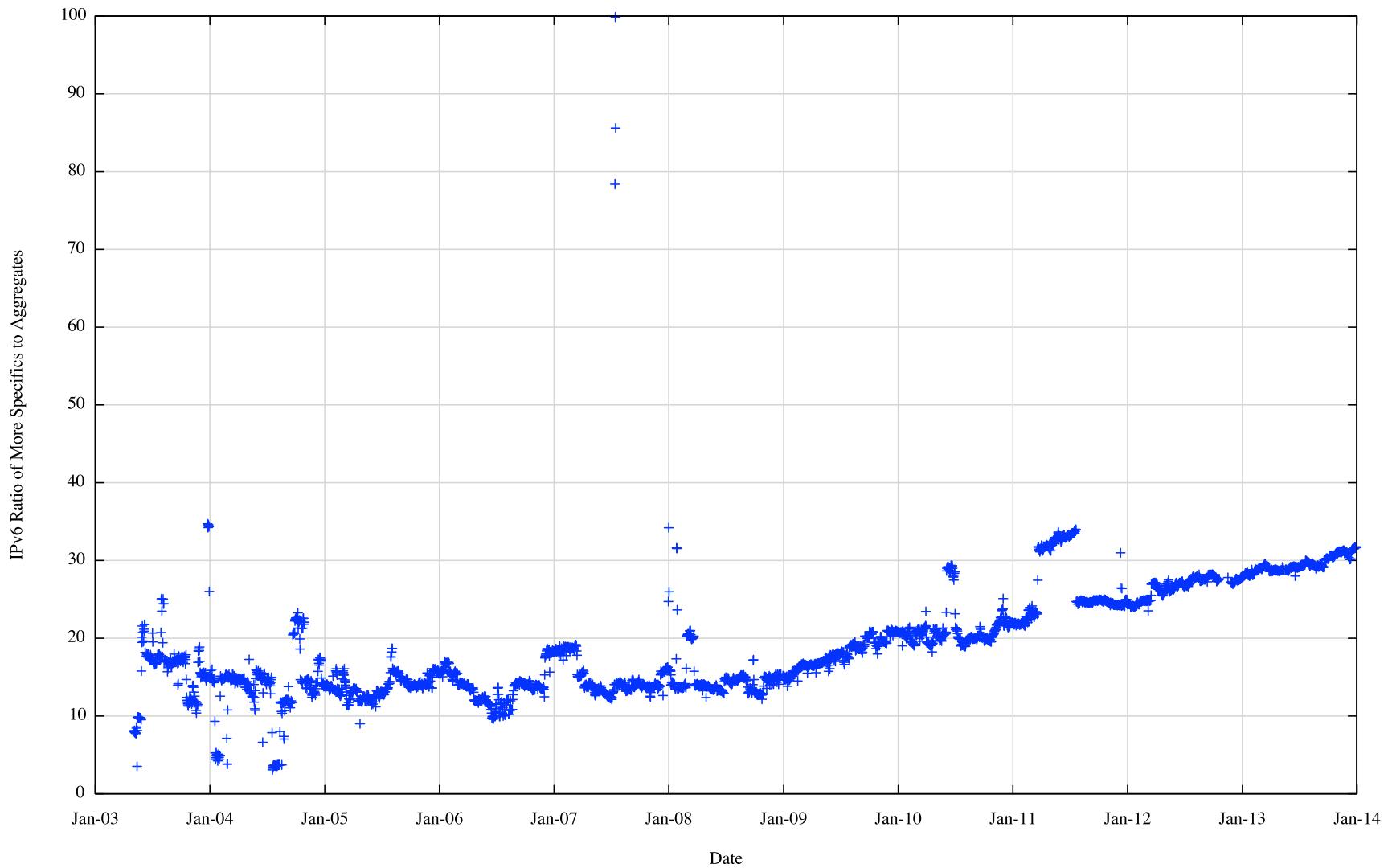
# Is it Everyone?

- 1% of the ASes (458 ASes) announce 54% of the more specifics (133,688 announcements)
- 55% of the ASes announce **no** more specifics
- The top 20 ASes announce 40,404 more specifics

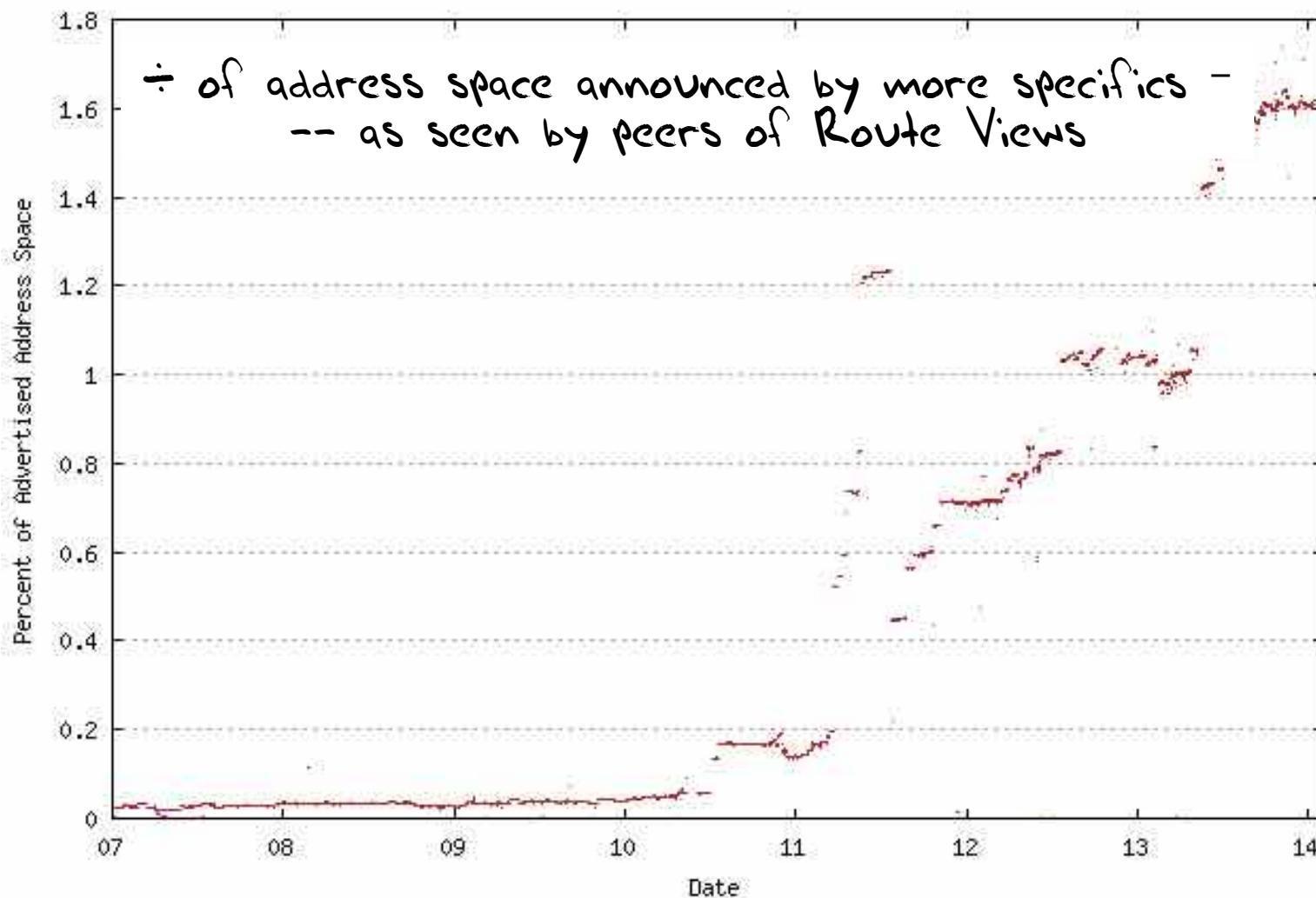
# More specifics in the V6 Routing Table



# More specifics in the V6 Routing Table



# How much V6 address space is announced by more specifics?



# Are We Getting Any Better?

Take the daily top 10 ASes of advertisers of more specifics over the past 3 years and track the number of more specifics advertised by these ASes over the entire period

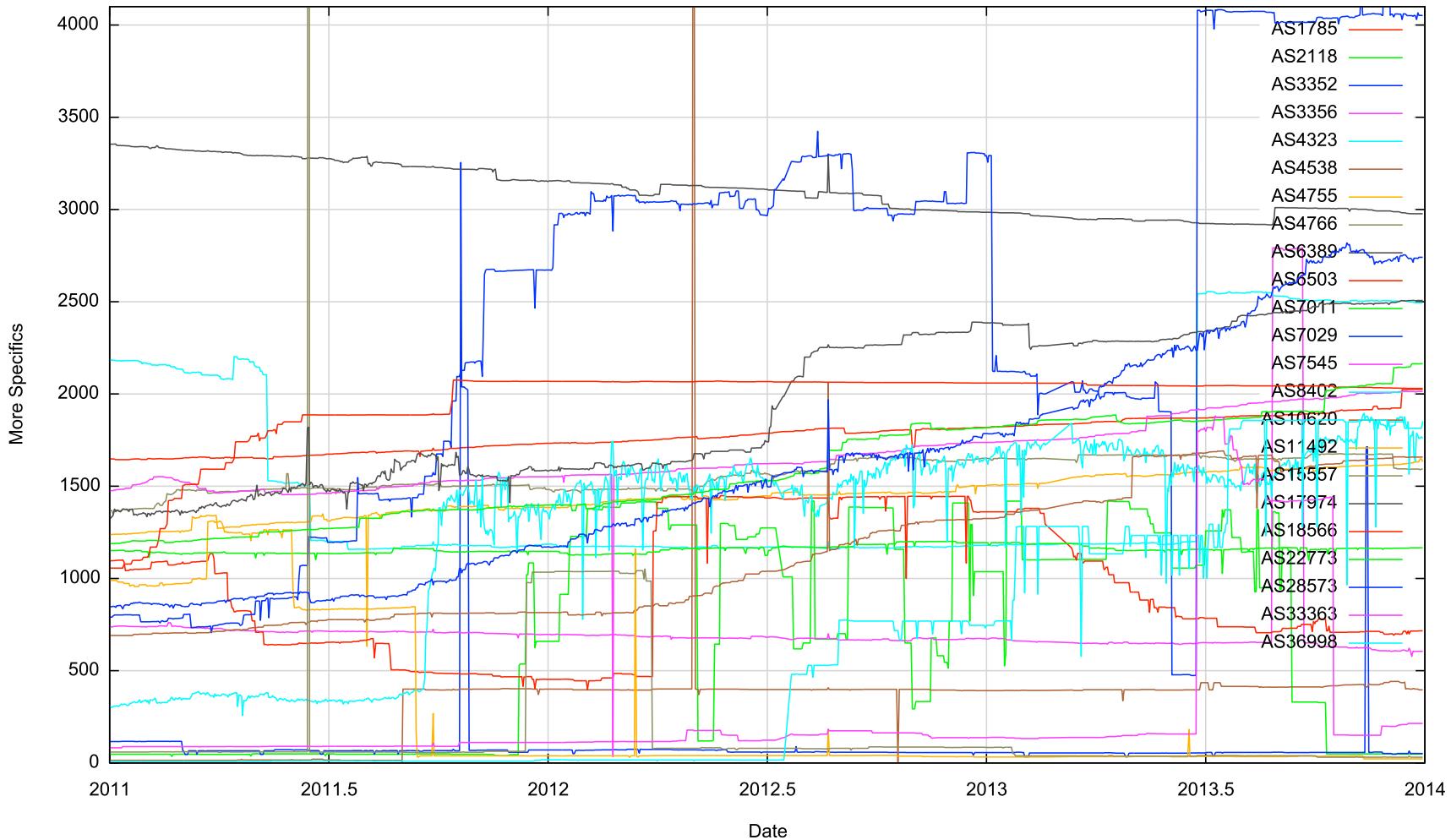
# Are We Getting any Better?

These AS's were seen to be advertising the highest number of more specifics over the past 3 years:

- 1785 AS-PAETEC-NET - PaeTec Communications, Inc. US
- 2118 RELCOM-AS OOO "NPO Relcom" RU
- 3352 TELEFONICA-DATA-ESPAÑA TELEFONICA DE ESPANA ES
- 3356 LEVEL3 Level 3 Communications US
- 4323 TWTC - tw telecom holdings, inc. US
- 4538 ERX-CERNET-BKB China Education and Research Network Center CN
- 4755 TATACOMM-AS TATA Communications formerly VSNL is Leading ISP IN
- 4766 KIXS-AS-KR Korea Telecom KR
- 6389 BELLSOUTH-NET-BLK - BellSouth.net Inc. US
- 6503 Axtel, S.A.B. de C.V. MX
- 7011 FRONTIER-AND-CITIZENS - Frontier Communications of America, Inc. US
- 7029 WINDSTREAM - Windstream Communications Inc US
- 7545 TPG-INTERNET-AP TPG Telecom Limited AU
- 8402 CORBINA-AS OJSC "Vimpelcom" RU
- 10620 Telmex Colombia S.A. CO
- 11492 CABLEONE - CABLE ONE, INC. US
- 15557 LDCOMNET Societe Francaise du Radiotelephone S.A FR
- 17974 TELKOMNET-AS2-AP PT Telekomunikasi Indonesia ID
- 18566 MEGAPATH5-US - MegaPath Corporation US
- 22773 ASN-CXA-ALL-CCI-22773-RDC - Cox Communications Inc. US
- 28573 NET Servicos de Comunicatio S.A. BR
- 33363 BHN-TAMPA - BRIGHT HOUSE NETWORKS, LLC US
- 36998 SDN-MOBITEL SD

# Yes ... and No

IPv4 More Specifics per AS: 2011 - 2013



# Are We Getting Any Better?

- Some ASes are effectively reducing the number of more specifics that are advertised into the global routing system
- Some ASes are increasing the number of more specifics
- And some are consistently advertising a significant number of more specifics
- There is no net change in the overall distribution and characteristics of more specifics in the routing system.

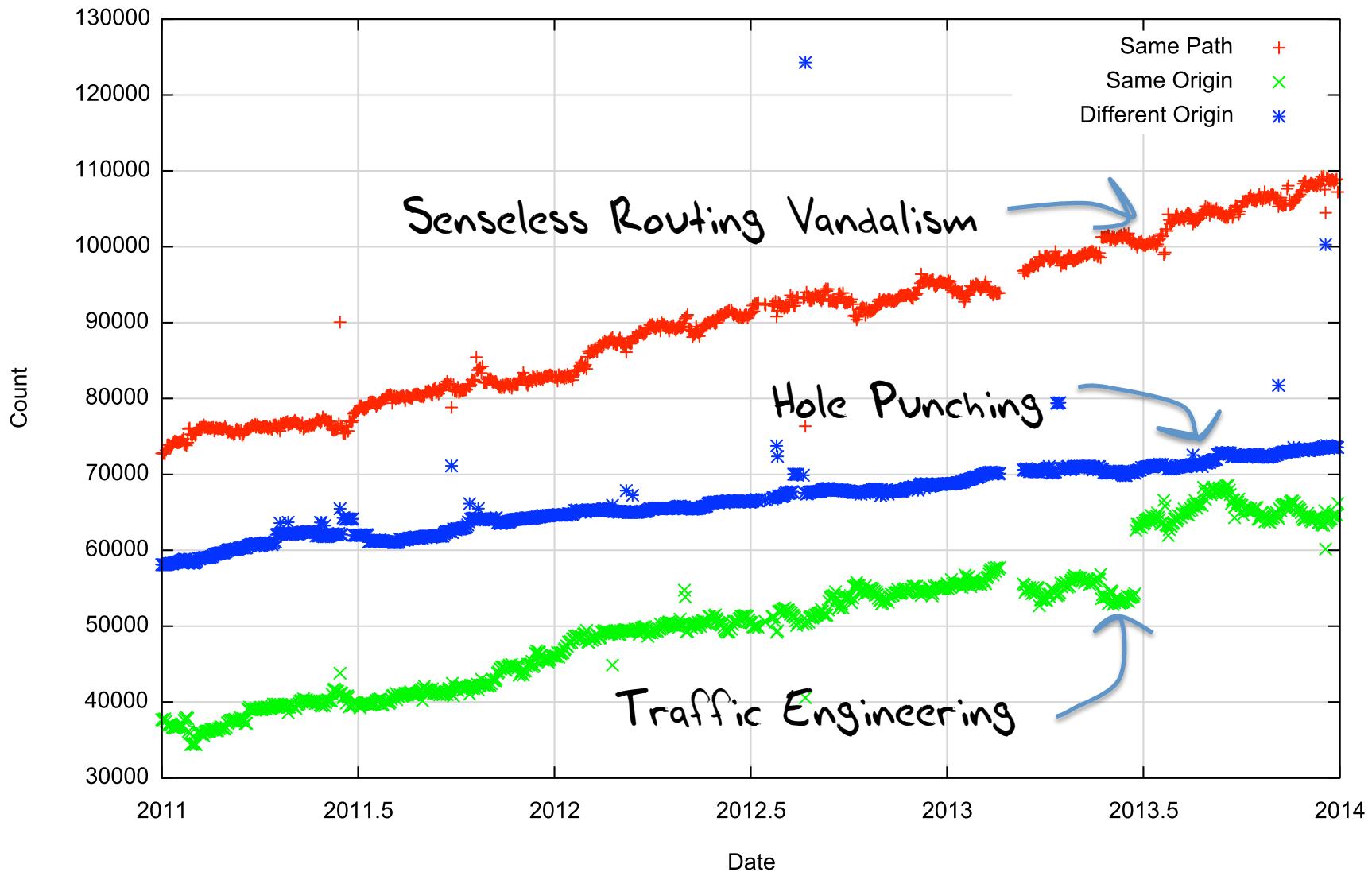
# Why advertise a more specific?

The reasons why we see more specifics in the routing system include:

- Different origination (“hole punching” in an aggregate)
- Traffic engineering of incoming traffic flows across multiple inter-AS paths
- “protection” against route hijacking by advertising more specifics
- Poor routing practices

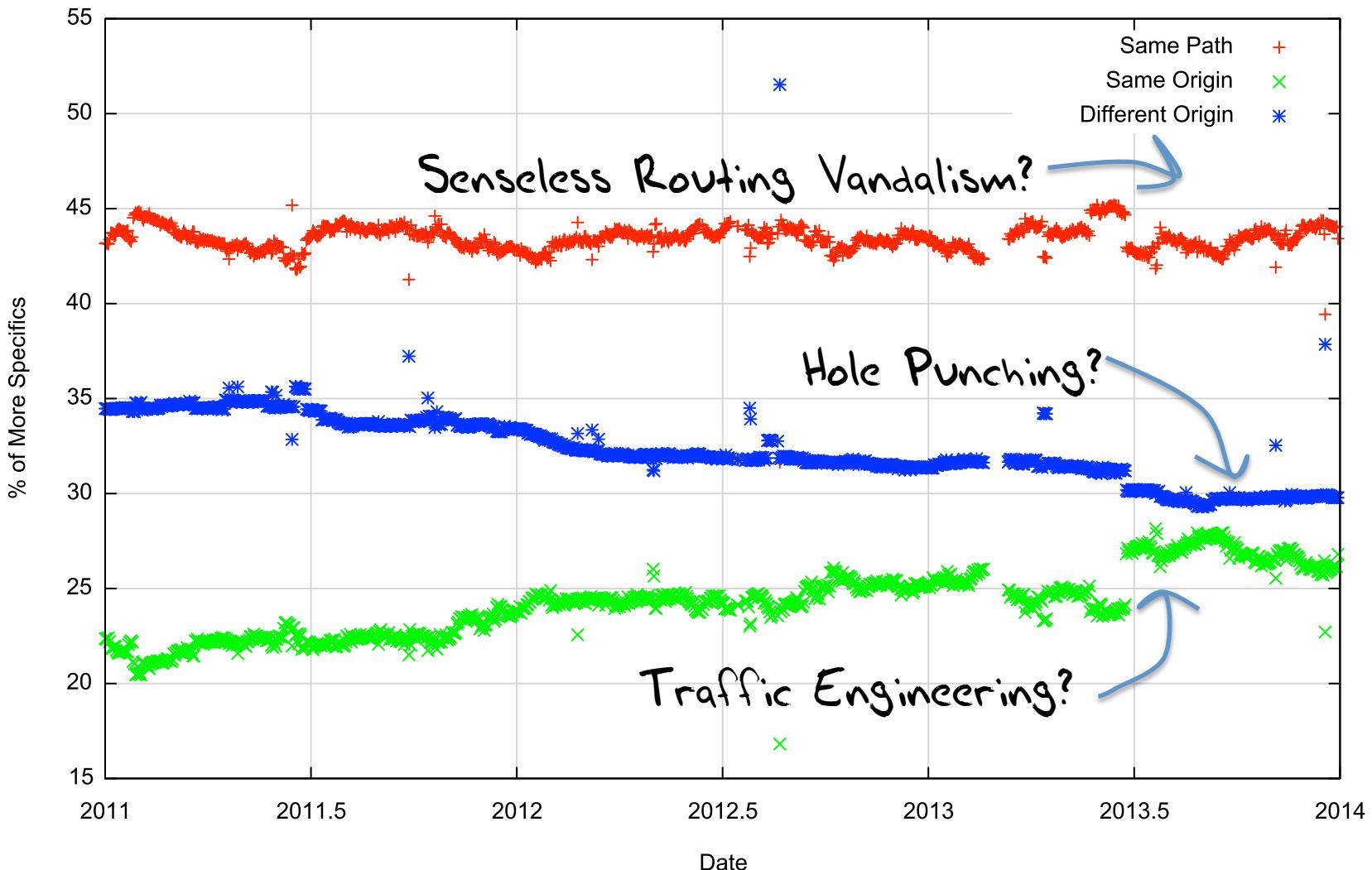
# Types of More Specifics

Type of More Specific



# Types of More Specifics

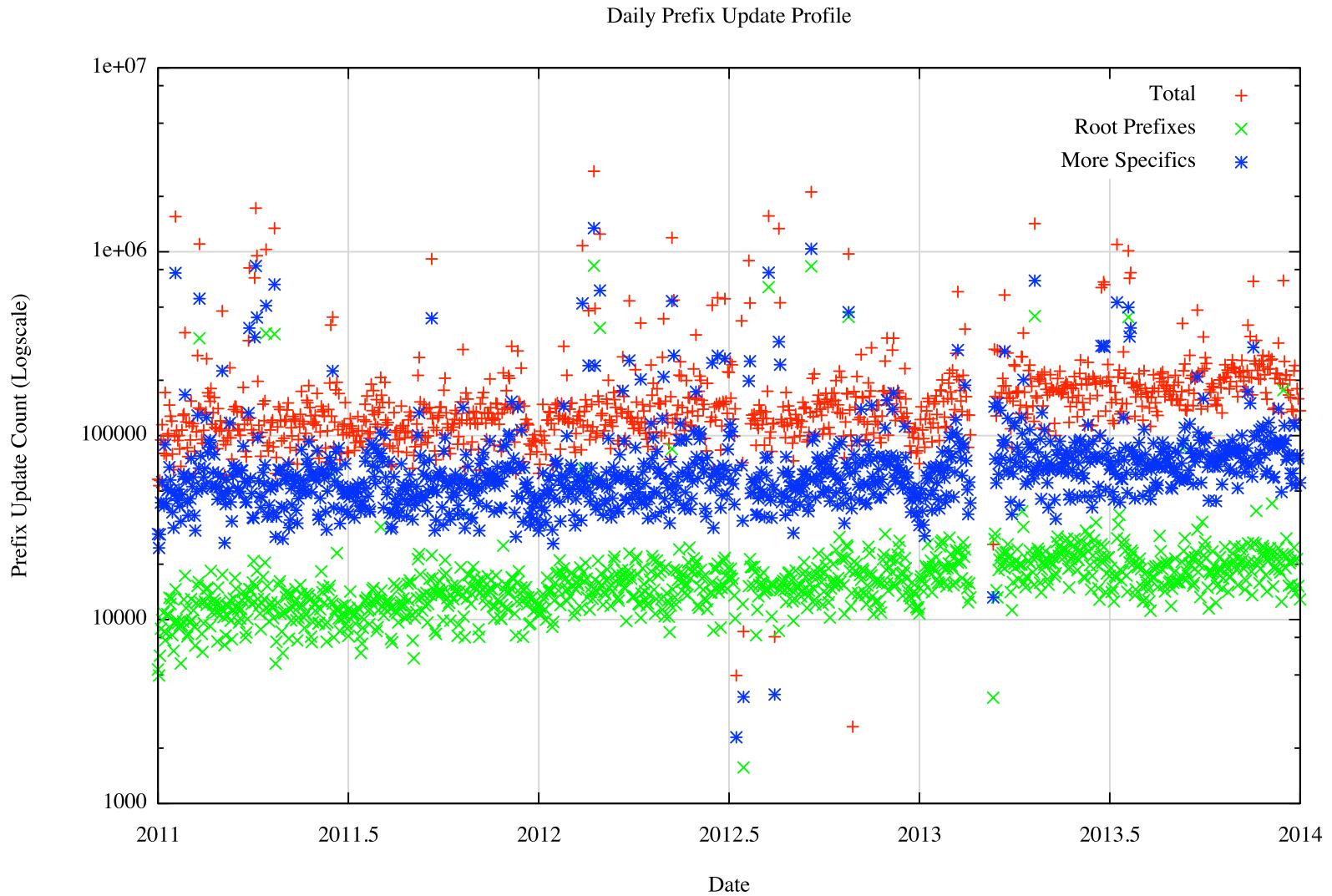
Relative Proportions of More Specifics



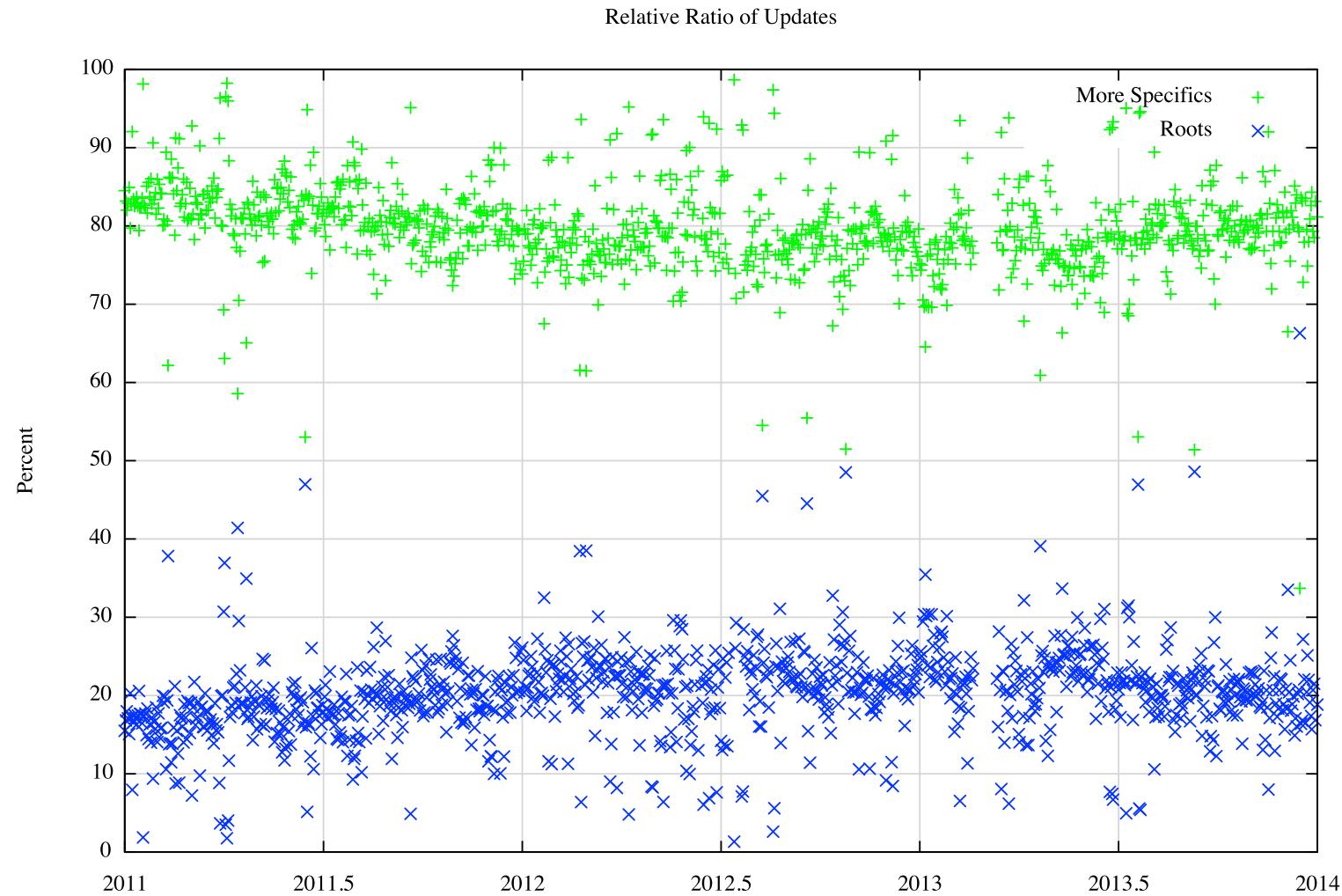
# Daily Update Rates

- Do more specifics experience a higher update rate than aggregate advertisements?
- Lets examine the past 3 years of updates and examine the daily count of prefix updates for root aggregates and more specifics

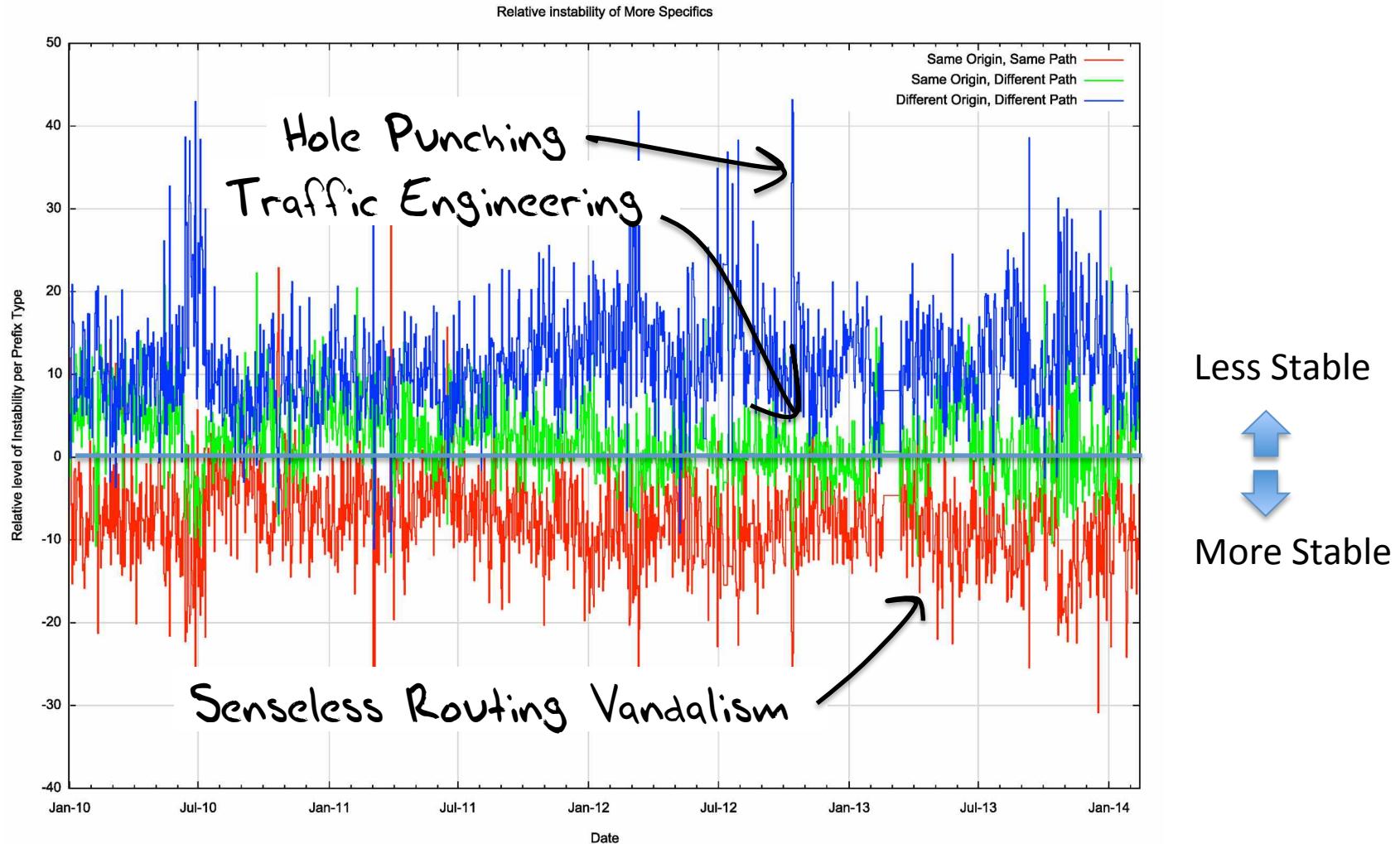
# Daily BGP Updates



# Relatively Speaking



# Relative Stability of More Specifics



# Daily Update Rates

- Do more specifics generate a higher update rate than aggregate advertisements?

Yes – in terms of prefix updates, more specifics are some 4 times noisier than the aggregates in terms of update traffic totals

More Specifics that “hole punch” (different origin AS) tend to be relatively noisier than other forms of more specifics

- Is this because they are less stable or noisier?

# What are we seeing?

- The profile of updates in BGP is dominated by the instability of the more specific announcements, which are 4 x more likely to experience instability compared to aggregate advertisements
- With the set of more specifics, “hole punching” (different origin AS, different AS Path) is consistently less stable than the other two types of more specifics.

# Problem? Not a Problem?

It's evident that the global BGP routing environment suffers from a certain amount of neglect and inattention

# Problem? Not a Problem?

It's evident that the global BGP routing environment suffers from a certain amount of neglect and inattention

*Could we do better?*

Yes!

Filtering out more specifics will both reduce table size and also reduce the level of dynamic update in the inter-domain environment

# Problem? Not a Problem?

It's evident that the global BGP routing environment suffers from a certain amount of neglect and inattention

*Should we do better?*

It can be difficult to justify the effort and the cost: the current growth rates of the routing table lie within relatively modest parameters of growth and still sit within the broad parameters of constant unit cost of routing technology

On the other hand, we need to recognize that we could do a lot better in terms of eliminating routing noise, and achieve this with a relatively modest amount of effort

Thank You

Questions?